1: J Invest Dermatol. 1999 Oct;113(4):643-50.

Regulation of epidermal bullous pemphigoid antigen 1 (BPAG1) synthesis by homeoprotein transcription factors.

Mainguy G, Erno H, Montesinos ML, Lesaffre B, Wurst W, Volovitch M, Prochiantz A.

CNRS, UMR 8542, Ecole Normale Superieure, Pans, France.

In a recent gene-trap screen, we identified the gene coding for Epidermal Bullous Pemphigoid Antigen 1 (BPAG1) as a putative transcriptional target of Engrailed and of other homeoproteins with a glutamine in position 50 of their homeodomain. We now show that the nuclear addressing of the homeodomains of Engrailed (EnHD) and Antennapedia (AntpHD) upregulates BPAG1e transcription in immortalized human keratinocytes (GMA24FIA) expressing En1. This upregulation is not observed with AntpHD-Q50A, a variant of AntpHD in which a single mutation abolishes its high-affinity binding to target DNA, thus strongly suggesting that BPAG1e upregulation homeodomains reflects their specific recognition of homeoprotein-binding sites in the BPAG1e locus. This is further confirmed by DNase I footprinting and electrophoretic mobility shift assays that reveal, within the cloned BPAG1e promoter, several sites of direct interaction with EnHD and Engrailed. Co-transfection experiments in GMA24FIA human keratinocytes, COS-7 simian fibroblasts, and CHP-100 human neuroepithelia cells show that Engrailed, Hoxa-5, and Hoxc-8 regulate BPAG1e promoter activity and that this regulation is context-dependent. Finally, using a mouse line with LacZ inserted within the En1 locus, we identify the keratinocytes of the ventral paws, including the epithelial cells of the eccrine tubules, as a strong site of En1 expression throughout adulthood. We therefore propose that BPAG1e, a 230 kDa keratin-binding protein expressed in keratinocytes and participating in the maintenance of hemidesmosomes at the dermis-epidermis border, is directly regulated by homeoprotein transcription factors.

PMID: 10504454 [PubMed - indexed for MEDLINE]

1: Mol Cells. 2000 Dec 31;10(6):728-32.

Efficient intracellular delivery of GFP by homeodomains of Drosophila Fushi-tarazu and Engrailed proteins.

Han K, Jeon MJ, Kim KA, Park J, Choi SY.

Department of Genetic Engineering, Hallym University, Chunchon, Korea.

The 60 amino acid long homeodomain of Antennapedia (Antp), either alone or as a fusion protein with 30-40 amino acid long foreign polypeptides, has been reported to cross biological membranes by an energy- and receptor-protein-independent mechanism. Moreover, the 16 amino acid long third helix of the Antp homeodomain, so-called penetratin, possesses translocation properties when fused to fewer than 100 amino acids as well. These findings led us to study whether such a protein tansduction property is shared by other homeodomains. We report here that homeodomains of two homeoproteins, Fushi-tarazu and Engrailed, are able to transduce a 238 amino acid long green fluorescent protein into cultured cells as efficiently as other well-known protein transduction domains, such as an internal oligopeptide of Tat and penetratin. These findings suggest that such transduction activity of homeodomains might have some physiological roles and that it can be exploited for development of efficient transduction vectors for research use and protein therapy.

PMID: 11211880 [PubMed - indexed for MEDLINE]

1: J Neurophysiol. 2000 Nov;84(5):2651-7.

Links

Topographical and physiological characterization of interneurons that express engrailed-1 in the embryonic chick spinal cord.

Wenner P, O'Donovan MJ, Matise MP.

National Institutes of Health, National Institute of Neurological Disorders and Stroke, Lab of Neural Control, Bethesda, Maryland 20892-4455, USA.

A number of homeodomain transcription factors have been implicated in controlling the differentiation of various types of neurons including spinal motoneurons. Some of these proteins are also expressed in spinal interneurons, but their function is unknown. Progress in understanding the role of transcription factors in interneuronal development has been slow because the synaptic connections of interneurons, which in part define their identity, are difficult to establish. Using whole cell recording in the isolated spinal cord of chick embryos, we assessed the synaptic connections of lumbosacral interneurons expressing the Engrailed-1 (En1) transcription factor. Specifically we established whether En1-expressing interneurons made direct connections with motoneurons and whether they constitute a single interneuron class. Cells were labeled with biocytin and subsequently processed for En1 immunoreactivity. Our findings indicate that the connections of En1-expressing cells with motoneurons and with sensory afferents were diverse, suggesting that the population was heterogeneous. In addition, the synaptic connections we tested were similar in interneurons that expressed the En1 protein and in many that did not. The majority of sampled En1 cells did, however, exhibit a direct synaptic connection to motoneurons that is likely to be GABAergic. Because our physiological methods underestimate the number of direct connections with motoneurons, it is possible that the great majority, perhaps all, En1-expressing cells make direct synaptic connections with motoneurons. Our results raise the possibility that En1 could be involved in interneuron-motoneuron connectivity but that its expression is not restricted to a distinct functional subclass of ventral interneuron. These findings constrain hypotheses about the role of En-1 in interneuron development and function.

PMID: 11068006 [PubMed - indexed for MEDLINE]

1: NP_001417. engrailed homolog...[gi:7710119] BLink, Domains, Links

```
LOCUS       NP_001417                392 aa            linear   PRI 04-OCT-2003
DEFINITION  engrailed homolog 1 [Homo sapiens].
ACCESSION   NP_001417
VERSION     NP_001417.2  GI:7710119
DBSOURCE    REFSEQ: accession NM_001426.2
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (residues 1 to 392)
  AUTHORS   Loomis,C.A., Harris,E., Michaud,J., Wurst,W., Hanks,M. and
            Joyner,A.L.
  TITLE     The mouse Engrailed-1 gene and ventral limb patterning
  JOURNAL   Nature 382 (6589), 360-363 (1996)
  MEDLINE   96300324
   PUBMED   8684466
REFERENCE   2  (residues 1 to 392)
  AUTHORS   Kohler,A., Logan,C., Joyner,A.L. and Muenke,M.
  TITLE     Regional assignment of the human homeobox-containing gene EN1 to
            chromosome 2q13-q21
  JOURNAL   Genomics 15 (1), 233-235 (1993)
  MEDLINE   93162666
   PUBMED   8094370
REFERENCE   3  (residues 1 to 392)
  AUTHORS   Logan,C., Hanks,M.C., Noble-Topham,S., Nallainathan,D.,
            Provart,N.J. and Joyner,A.L.
  TITLE     Cloning and sequence comparison of the mouse, human, and chicken
            engrailed genes reveal potential functional domains and regulatory
            regions
  JOURNAL   Dev. Genet. 13 (5), 345-358 (1992)
  MEDLINE   93185339
   PUBMED   1363401
REFERENCE   4  (residues 1 to 392)
  AUTHORS   Logan,C., Willard,H.F., Rommens,J.M. and Joyner,A.L.
  TITLE     Chromosomal localization of the human homeo box-containing genes,
            EN1 and EN2
  JOURNAL   Genomics 4 (2), 206-209 (1989)
  MEDLINE   89290849
   PUBMED   2567700
COMMENT     REVIEWED REFSEQ: This record has been curated by NCBI staff. The
            reference sequence was derived from L12699.1.
            On May 5, 2000 this sequence version replaced gi:4503567.

            Summary: Homeobox-containing genes are thought to have a role in
            controlling development. In Drosophila, the 'engrailed' (en) gene
            plays an important role during development in segmentation, where
```

```
            it is required for the formation of posterior compartments.
            Different mutations in the mouse homologs, En1 and En2, produced
            different developmental defects that frequently are lethal. The
            human engrailed homologs 1 and 2 encode homeodomain-containing
            proteins and have been implicated in the control of pattern
            formation during development of the central nervous system.
FEATURES             Location/Qualifiers
     source          1..392
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2q13-q21"
     Protein         1..392
                     /product="engrailed homolog 1"
     Region          303..362
                     /region_name="homeobox domain"
     Region          305..360
                     /region_name="Homeobox domain"
                     /note="homeobox"
                     /db_xref="CDD:pfam00046"
     CDS             1..392
                     /gene="EN1"
                     /coded_by="NM_001426.2:1016..2194"
                     /note="go_component: nucleus [goid 0005634] [evidence
                     IEA];
                     go_function: transcription factor activity [goid 0003700]
                     [evidence IEA];
                     go_process: skeletal development [goid 0001501] [evidence
                     TAS] [pmid 8684466];
                     go_process: embryogenesis and morphogenesis [goid 0007345]
                     [evidence TAS] [pmid 8094370];
                     go_process: regulation of transcription, DNA-dependent
                     [goid 0006355] [evidence IEA];
                     go_process: development [goid 0007275] [evidence IEA]"
                     /db_xref="GeneID:2019"
                     /db_xref="LocusID:2019"
                     /db_xref="MIM:131290"
ORIGIN
        1 meeqqpepks qrdsalggaa aatpgglsls lspgasgssg sgsdgdsvpv spqpappspp
       61 aapclpplah hphlpphppp pppqhlaapa hqpqpaaqlh rttnffidni lrpdfgckke
      121 qpppqllvaa aarggagggg rverdrgqta agrdpvhplg trapgaasll capdancgpp
      181 dgsqpaaaga gaskagnpaa aaaaaaaava aaaaaaaakp sdtggggsgg gagspgaqgt
      241 kypehgnpai llmgsanggp vvktdsqqpl vwpawvyctr ysdrpssgpr trklkkkkne
      301 kedkrprtaf taeqlqrlka efqanryite qrrqtlaqel slnesqikiw fqnkrakikk
      361 atgiknglal hlmaqglynh stttvqdkde se
//
```

1: A48423. engrailed homeodo...[gi:477223] BLink, Domains, Links

```
LOCUS       A48423                  401 aa            linear   ROD 24-SEP-1999
DEFINITION  engrailed homeodomain-containing protein En-1 - mouse.
ACCESSION   A48423
VERSION     A48423  GI:477223
DBSOURCE    pir: locus A48423;

            summary: #length 401 #molecular-weight 40950 #checksum 1310
            ;
            genetic: #gene en.1 #map_position 1
            ;
            superfamily: unassigned homeobox proteins; homeobox homology
            ;
            PIR dates: 01-Dec-1993 #sequence_revision 18-Nov-1994 #text_change
            24-Sep-1999
            .
KEYWORDS    DNA binding; homeobox; nucleus; transcription regulation.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (residues 1 to 401)
  AUTHORS   Joyner,A.L., Kornberg,T., Coleman,K.G., Cox,D.R. and Martin,G.R.
  TITLE     Expression during embryogenesis of a mouse gene with sequence
            homology to the Drosophila engrailed gene
  JOURNAL   Cell 43 (1), 29-37 (1985)
  MEDLINE   86079501
   PUBMED   2416459
REFERENCE   2  (residues 1 to 401)
  AUTHORS   Joyner,A.L. and Martin,G.R.
  TITLE     En-1 and En-2, two mouse genes with sequence homology to the
            Drosophila engrailed gene: expression during embryogenesis
  JOURNAL   Genes Dev. 1 (1), 29-38 (1987)
  MEDLINE   88112776
   PUBMED   2892757
REFERENCE   3  (residues 1 to 401)
  AUTHORS   Holland,P.W. and Williams,N.A.
  TITLE     Conservation of engrailed-like homeobox sequences during vertebrate
            evolution
  JOURNAL   FEBS Lett. 277 (1-2), 250-252 (1990)
  MEDLINE   91099509
   PUBMED   1980115
REFERENCE   4  (residues 1 to 401)
  AUTHORS   Logan,C., Hanks,M.C., Noble-Topham,S., Nallainathan,D.,
            Provart,N.J. and Joyner,A.L.
  TITLE     Cloning and sequence comparison of the mouse, human, and chicken
            engrailed genes reveal potential functional domains and regulatory
            regions
```

```
  JOURNAL   Dev. Genet. 13 (5), 345-358 (1992)
  MEDLINE   93185339
   PUBMED   1363401
FEATURES             Location/Qualifiers
     source          1..401
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     Protein         1..401
                     /product="engrailed homeodomain-containing protein En-1"
                     /note="homeotic protein En-1"
     Region          313..369
                     /region_name="domain"
                     /note="homeobox homology #label HOX"
ORIGIN
        1 meeqqpepks qrdsglgava aaapsglsls lspgasgssg sdgdsvpvsp qpappsppaa
       61 pclpplahhp hlpphppppp pppppppqhl aapahqpqpa aqlhrttnff idnilrpdfg
      121 ckkeqplpql lvasaaaggg aaagggsrve rdrgqtgagr dpvhslgtra sgaasllcap
      181 dancgppdgs qpatavgaga skagnpaaaa aaaaaaaaaa vaaaaaaask psdsgggsgg
      241 nagspgaqga kfpehnpail lmgsanggpv vktdsqqplv wpawvyctry sdrpssgprt
      301 rklkkkknek edkrprtaft aeqlqrlkae fqanryiteq rrqtlaqels lnesqikiwf
      361 qnkrakikka tgiknglalh lmaqglynhs tttvqdkdes e
//
```


Nov 18 2003 07:36:37

Entrez Protein

Entrez PubMed Nucleotide Protein Genome Structure PMC Taxonomy Boo

Search Protein for Go Clear

Limits Preview/Index History Clipboard Details

Display default Show: 20 Send to File Get Subsequence Featu

1: B25682. homeotic protein ...[gi:85248] BLink, Domains, Links

```
LOCUS       B25682                  584 aa            linear   INV 20-AUG-1999
DEFINITION  homeotic protein Engrailed - fruit fly (Drosophila virilis).
ACCESSION   B25682
VERSION     B25682  GI:85248
DBSOURCE    pir: locus B25682;

            summary: #length 584 #molecular-weight 62177 #checksum 3887
            ;
            genetic: #gene en ##cross-references FlyBase:FBgn0013111 #introns
            470/1; 502/3
            ;
            superfamily: engrailed homeotic protein; homeobox homology
            ;
            PIR dates: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change
            20-Aug-1999
            .
KEYWORDS    DNA binding; homeobox; nucleus; transcription regulation.
SOURCE      Drosophila virilis
  ORGANISM  Drosophila virilis
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (residues 1 to 584)
  AUTHORS   Kassis,J.A., Poole,S.J., Wright,D.K. and O'Farrell,P.H.
  TITLE     Sequence conservation in the protein coding and intron regions of
            the engrailed transcription unit
  JOURNAL   EMBO J. 5 (13), 3583-3589 (1986)
  MEDLINE   87161768
   PUBMED   2881781
FEATURES             Location/Qualifiers
     source          1..584
                     /organism="Drosophila virilis"
                     /db_xref="taxon:7244"
     Protein         1..584
                     /product="homeotic protein Engrailed"
     Region          487..543
                     /region_name="domain"
                     /note="homeobox homology #label HOX"
ORIGIN
        1 maledrcspq sapsppgclp hsppqqhqhi lqpnlrihai pisitaldms lqpgtaataa
       61 tlyqqqqlqh lhqlqqlqql hqqqlaasag vfhhpstsaa fdaaalhaaa lqqrlsasgs
      121 pgahstgacs tpastltike eesdsilgds fhnqtattat eedeeedddi dvddtasvsg
      181 sarlpppahq qskqskpsla fsisnilsdr fgdaakqtkq gdstgvsvsa aaaaasifrp
      241 feasraaaat psaftrvdll efsrqqqaaa aaataammle ranllncfnp aaypriheel
      301 vqsrlrrsaa gnpnavlppt akltepsvek salgslcktv sqigqaqstt pvttpssrps
      361 qlaspppasn astissssst aascsassss gcssaassln sspssrlaga nnasspqpip
      421 ppsavsrdsg messddtrse tgstttdggk nemwpawvyc trysdrpssg pryrrpkqpk
```

```
      481 dktndekrpr tafsseqlar lkrefnenry lterrrqqls selglneaqi kiwfqnkrak
      541 ikkstgsknp lalqlmaqgl ynhttvpltk eeeelemrmn gqip
//
```

1: P02833. Homeotic antennap...[gi:123317] BLink, Domains, Links

```
LOCUS       P02833                    378 aa            linear   INV 15-SEP-2003
DEFINITION  Homeotic antennapedia protein.
ACCESSION   P02833
VERSION     P02833  GI:123317
DBSOURCE    swissprot: locus HMAN_DROME, accession P02833;
            class: standard.
            extra accessions:Q95SZ6,created: Jul 21, 1986.
            sequence updated: Nov 1, 1988.
            annotation updated: Sep 15, 2003.
            xrefs: gi: 7593, gi: 1805742, gi: 7595, gi: 156948, gi: 156949, gi:
            156950, gi: 156951, gi: 14288990, gi: 4389423, gi: 23170613, gi:
            28381157, gi: 28381155, gi: 16648361, gi: 16648362, gi: 156945, gi:
            156947, gi: 156931, gi: 156934, gi: 156939, gi: 156940, gi: 84890,
            pdb accession 1AHD, gi: 443020, gi: 1431670, gi: 1421329, pdb
            accession 9ANT
            xrefs (non-sequence databases): TRANSFACT00026, FlyBaseFBgn0000095,
            InterProIPR001827, InterProIPR001356, PfamPF00046, PRINTSPR00025,
            PRINTSPR00024, ProDomPD000010, SMARTSM00389, PROSITEPS00032,
            PROSITEPS00027, PROSITEPS50071
KEYWORDS    Homeobox; DNA-binding; Developmental protein; Nuclear protein;
            Repeat; Alternative splicing; 3D-structure.
SOURCE      Drosophila melanogaster (fruit fly)
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (residues 1 to 378)
  AUTHORS   Schneuwly,S., Kuroiwa,A., Baumgartner,P. and Gehring,W.J.
  TITLE     Structural organization and sequence of the homeotic gene
            Antennapedia of Drosophila melanogaster
  JOURNAL   EMBO J. 5 (4), 733-739 (1986)
  MEDLINE   99334708
   PUBMED   10408949
  REMARK    SEQUENCE FROM N.A. (ISOFORM 1).
REFERENCE   2  (residues 1 to 378)
  AUTHORS   Stroeher,V.L., Jorgensen,E.M. and Garber,R.L.
  TITLE     Multiple transcripts from the Antennapedia gene of Drosophila
            melanogaster
  JOURNAL   Mol. Cell. Biol. 6 (12), 4667-4675 (1986)
  MEDLINE   87089828
   PUBMED   2879222
  REMARK    SEQUENCE FROM N.A. (ISOFORM 1).
            STRAIN=Oregon-R; TISSUE=Embryo, Larva, and Pupae
REFERENCE   3  (residues 1 to 378)
  AUTHORS   Laughon,A., Boulet,A.M., Bermingham,J.R. Jr., Laymon,R.A. and
            Scott,M.P.
  TITLE     Structure of transcripts from the homeotic Antennapedia gene of
```

Drosophila melanogaster: two promoters control the major protein-coding region
JOURNAL Mol. Cell. Biol. 6 (12), 4676-4689 (1986)
MEDLINE 87089829
PUBMED 2879223
REMARK SEQUENCE FROM N.A. (ISOFORM 1).
REFERENCE 4 (residues 1 to 378)
AUTHORS Celniker,S.E., Pfeiffer,B., Knafels,J., Martin,C.H., Mayeda,C.A. and Palazzolo,M.J.
TITLE Direct Submission
JOURNAL Submitted (~JAN-1999)
REMARK SEQUENCE FROM N.A. (ISOFORM 1).
REFERENCE 5 (residues 1 to 378)
AUTHORS Adams,M.D., Celniker,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D., Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galle,R.F., George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N., Sutton,G.G., Wortman,J.R., Yandell,M.D., Zhang,Q., Chen,L.X., Brandon,R.C., Rogers,Y.-H.C., Blazej,R.G., Champe,M., Pfeiffer,B.D., Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Miklos,G.L.G., Abril,J.F., Agbayani,A., An,H.-J., Andrews-Pfannkoch,C., Baldwin,D., Ballew,R.M., Basu,A., Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y., Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D., Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C., Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I., Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M., Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C., Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferriera,S., Fleischmann,W., Fosler,C., Gabrielian,A.E., Garg,N.S., Gelbart,W.M., Glasser,K., Glodek,A., Gong,F., Gorrell,J.H., Gu,Z., Guan,P., Harris,M., Harris,N.L., Harvey,D.A., Heiman,T.J., Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J., Wei,M.-H., Ibegwam,C., Jalali,M., Kalush,F., Karpen,G.H., Ke,Z., Kennison,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C., Kravitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A., Li,J.H., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,B., McIntosh,T.C., McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C., Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L., Muzny,D.M., Nelson,D.L., Nelson,D.R., Nelson,K.A., Nixon,K., Nusskern,D.R., Pacleb,J.M., Palazzolo,M., Pittman,G.S., Pan,S., Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K., Saunders,R.D.C., Scheeler,F., Shen,H., Shue,B.C., Siden-Kiamos,I., Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C., Stapleton,M., Strong,R., Sun,E., Svirskas,R., Tector,C., Turner,R., Venter,E., Wang,A.H., Wang,X., Wang,Z.-Y., Wassarman,D.A., Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T., Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.-F., Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H., Zhong,F.N., Zhong,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O., Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
TITLE The genome sequence of Drosophila melanogaster
JOURNAL Science 287 (5461), 2185-2195 (2000)
MEDLINE 20196006
PUBMED 10731132
REMARK SEQUENCE FROM N.A.
STRAIN=Berkeley
REFERENCE 6 (residues 1 to 378)
AUTHORS Misra,S., Crosby,M.A., Mungall,C.J., Matthews,B.B., Campbell,K.S., Hradecky,P., Huang,Y., Kaminker,J.S., Millburn,G.H., Prochnik,S.E.,

Smith,C.D., Tupy,J.L., Whitfield,E.J., Bayraktaroglu,L., Berman,B.P., Bettencourt,B.R., Celniker,S.E., de Grey,A.D.N.J., Drysdale,R.A., Harris,N.L., Richter,J., Russo,S., Schroeder,A.J., Shu,S.Q., Stapleton,M., Yamada,C., Ashburner,M., Gelbart,W.M., Rubin,G.M. and Lewis,S.E.
TITLE Annotation of the Drosophila melanogaster euchromatic genome: a systematic review
JOURNAL Genome Biol. 3 (12), RESEARCH0083 (2002)
MEDLINE 22426069
PUBMED 12537572
REMARK REVISIONS, AND ALTERNATIVE SPLICING.
REFERENCE 7 (residues 1 to 378)
AUTHORS Stapleton,M., Carlson,J.W., Brokstein,P., Yu,C., Champe,M., George,R.A., Guarin,H., Kronmiller,B., Pacleb,J.M., Park,S., Wan,K.H., Rubin,G.M. and Celniker,S.E.
TITLE A Drosophila full-length cDNA resource
JOURNAL Genome Biol. 3 (12), RESEARCH0080 (2002)
MEDLINE 22426066
PUBMED 12537569
REMARK SEQUENCE FROM N.A. (ISOFORM 2).
STRAIN=Berkeley; TISSUE=Embryo
REFERENCE 8 (residues 1 to 378)
AUTHORS Scott,M.P. and Weiner,A.J.
TITLE Structural relationships among genes that control development: sequence homology between the Antennapedia, Ultrabithorax, and fushi tarazu loci of Drosophila
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (13), 4115-4119 (1984)
MEDLINE 84248068
PUBMED 6330741
REMARK SEQUENCE OF 296-364 FROM N.A.
REFERENCE 9 (residues 1 to 378)
AUTHORS McGinnis,W., Garber,R.L., Wirz,J., Kuroiwa,A. and Gehring,W.J.
TITLE A homologous protein-coding sequence in Drosophila homeotic genes and its conservation in other metazoans
JOURNAL Cell 37 (2), 403-408 (1984)
MEDLINE 84205674
PUBMED 6327065
REMARK SEQUENCE OF 296-378 FROM N.A.
REFERENCE 10 (residues 1 to 378)
AUTHORS Regulski,M., Harding,K., Kostriken,R., Karch,F., Levine,M. and McGinnis,W.
TITLE Homeo box genes of the Antennapedia and bithorax complexes of Drosophila
JOURNAL Cell 43 (1), 71-80 (1985)
MEDLINE 86079516
PUBMED 2416463
REMARK SEQUENCE OF 297-357 FROM N.A.
REFERENCE 11 (residues 1 to 378)
AUTHORS Schneuwly,S., Klemenz,R. and Gehring,W.J.
TITLE Redesigning the body plan of Drosophila by ectopic expression of the homoeotic gene Antennapedia
JOURNAL Nature 325 (6107), 816-818 (1987)
MEDLINE 87144589
PUBMED 3821869
REMARK MUTANT ANALYSIS.
REFERENCE 12 (residues 1 to 378)
AUTHORS Billeter,M., Qian,Y., Otting,G., Muller,M., Gehring,W.J. and Wuthrich,K.
TITLE Determination of the three-dimensional structure of the Antennapedia homeodomain from Drosophila in solution by 1H nuclear

PUBMED 2104909
REMARK STRUCTURE BY NMR OF HOMEOBOX.
REFERENCE 13 (residues 1 to 378)
AUTHORS Qian,Y.Q., Otting,G., Billeter,M., Muller,M., Gehring,W. and Wuthrich,K.
TITLE Nuclear magnetic resonance spectroscopy of a DNA complex with the uniformly 13C-labeled Antennapedia homeodomain and structure determination of the DNA-bound homeodomain
JOURNAL J. Mol. Biol. 234 (4), 1070-1083 (1993)
MEDLINE 94087721
PUBMED 7903397
REMARK STRUCTURE BY NMR OF HOMEOBOX.
REFERENCE 14 (residues 1 to 378)
AUTHORS Billeter,M., Qian,Y.Q., Otting,G., Muller,M., Gehring,W. and Wuthrich,K.
TITLE Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex
JOURNAL J. Mol. Biol. 234 (4), 1084-1093 (1993)
MEDLINE 94087722
PUBMED 7903398
REMARK STRUCTURE BY NMR OF HOMEOBOX.
REFERENCE 15 (residues 1 to 378)
AUTHORS Qian,Y.Q., Otting,G., Furukubo-Tokunaga,K., Affolter,M., Gehring,W.J. and Wuthrich,K.
TITLE NMR structure determination reveals that the homeodomain is connected through a flexible linker to the main body in the Drosophila Antennapedia protein
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 89 (22), 10738-10742 (1992)
MEDLINE 93066318
PUBMED 1359544
REMARK STRUCTURE BY NMR OF 279-363.
REFERENCE 16 (residues 1 to 378)
AUTHORS Fraenkel,E. and Pabo,C.O.
TITLE Comparison of X-ray and NMR structures for the Antennapedia homeodomain-DNA complex
JOURNAL Nat. Struct. Biol. 5 (8), 692-697 (1998)
MEDLINE 98363212
PUBMED 9699632
REMARK X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 297-356.
COMMENT ----------

[FUNCTION] Sequence-specific transcription factor which is part of a developmental regulatory system that regulates segmental identity in the mesothorax. Provides cells with specific positional identities on the anterior-posterior axis.
[SUBCELLULAR LOCATION] Nuclear (Probable).
[ALTERNATIVE PRODUCTS] Event=Alternative splicing; Named isoforms=2; Name=1; Synonyms=A, B, I, J; IsoId=P02833-1; Sequence=Displayed; Name=2; Synonyms=H; IsoId=P02833-2; Sequence=VSP_008097, VSP_008098; Note=No experimental confirmation available.
[MISCELLANEOUS] Loss of Antp results in altered development of the

```
            [CAUTION] Ref.5 (AAN13333) sequence differs from that shown due to
            erroneous gene model prediction.
FEATURES             Location/Qualifiers
     source          1..378
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
     gene            1..378
                     /gene="ANTP"
                     /note="synonyms: BG:DS07700.1, CG1028"
     Protein         1..378
                     /gene="ANTP"
                     /product="Homeotic antennapedia protein"
     Region          76..81
                     /gene="ANTP"
                     /region_name="Domain"
                     /note="GLN-RICH (OPA-REPEAT)."
     Region          110..155
                     /gene="ANTP"
                     /region_name="Domain"
                     /note="GLN-RICH (OPA-REPEAT)."
     Site            283..288
                     /gene="ANTP"
                     /site_type="unclassified"
                     /note="ANTP-TYPE HEXAPEPTIDE."
     Region          296..297
                     /gene="ANTP"
                     /region_name="Splicing variant"
                     /note="ER -> GK (in isoform 2). /FTId=VSP_008097."
     Site            297..356
                     /gene="ANTP"
                     /site_type="DNA binding"
                     /note="HOMEOBOX."
     Region          298..378
                     /gene="ANTP"
                     /region_name="Splicing variant"
                     /note="Missing (in isoform 2). /FTId=VSP_008098."
     Region          300
                     /gene="ANTP"
                     /region_name="Conflict"
                     /note="G -> E (IN REF. 10)."
     Region          306..316
                     /gene="ANTP"
                     /region_name="Helical region"
     Region          317..318
                     /gene="ANTP"
                     /region_name="Hydrogen bonded turn"
     Region          324..334
                     /gene="ANTP"
                     /region_name="Helical region"
     Region          335
                     /gene="ANTP"
                     /region_name="Hydrogen bonded turn"
     Region          338..354
                     /gene="ANTP"
                     /region_name="Helical region"
ORIGIN
        1 mtmstnnces mtsyftnsym gadmhhghyp gngvtdldaq qmhhysqnan hqgnmpyprf
```

241 gappqqqmmmrq qqqppqmmqqy mpqqmtppsq mpmsqssgmp spiypwmmrsq rgkcqerkrg
301 rqtytryqtl elekefhfnr yltrrrriei ahalclterq ikiwfqnrrm kwkkenktkg
361 epgsggegde itppnspq
//

Nov 18 2003 07:36:37

COMMUNICATION

ChemComm
www.rsc.org/chemcomm

Gadolinium-binding helix–turn–helix peptides: DNA-dependent MRI contrast agents†

Peter Caravan,[a] Jaclyn M. Greenwood,[a] Joel T. Welch[b] and Sonya J. Franklin*[b]

[a] *EPIX Medical, Inc., 71 Rogers Street, Cambridge, MA 02142, USA*

[b] *Department of Chemistry, University of Iowa, Iowa City, IA 52242, USA. E-mail: sonya-franklin@uiowa.edu; Fax: 01 319-353-2244; Tel: 01 319-335-1270*

Received (in Cambridge, MA, USA) 9th July 2003, Accepted 12th August 2003
First published as an Advance Article on the web 12th September 2003

A *de novo* designed gadolinium metallopeptide was found to be a very efficient relaxation agent, with 100% increase in relaxivity upon binding to DNA.

Molecular imaging has been defined as "the *in vivo* characterization and measurement of biologic processes at the cellular and molecular level."[1] Magnetic resonance imaging (MRI) has great potential to play a broader role in this field, if the poor sensitivity of MRI contrast agents can be overcome. A molecular imaging MRI contrast agent must therefore combine a targeting vector (recognizing a specific biomarker) with a high relaxivity entity (a paramagnetic group that makes the interaction detectable by MR). Several different approaches to high relaxivity have been reported, including targeted superparamagnetic particles,[2] targeted gadolinium-containing emulsions[3] or complex assemblies,[4] and enzyme amplified contrast strategies.[5] This report focuses on designing a high relaxivity Gd(III) chelate within a targeting moiety, and further amplifying the relaxivity by binding to a receptor (DNA).

The receptor induced magnetization enhancement (RIME) effect is an effective MRI signal amplification strategy,[6–8] in which receptor binding increases the reorientation time (τ_R), thereby increasing the relaxation enhancement of solvent proton relaxation rates ($1/T_1$). Targeted MRI contrast agents that utilize the RIME effect are typically bifunctional molecules: a Gd(III) chelate complex appended to a targeting moiety through a tether (Fig. 1). However, the linker group may introduce rotational flexibility and limit the gain in relaxivity ($r_1 = (\Delta 1/T_1)/[Gd]$) brought about by receptor binding. If the chelating moiety could be incorporated within the targeting group, internal motion and rotational flexibility may be reduced and greater relaxivities could be achieved upon binding.

Peptide-based chelators are an attractive platform for incorporating a Gd site into a biological recognition unit. Substitution of Gd into Ca binding proteins (O_8 or O_9 donor set) can yield very high relaxivities.[9] A peptide-based Gd-chelator mimicking a Ca binding protein thus offers the potential of high relaxivity and a means to exploit protein structure–function relationships in recognition. Solid phase peptide synthesis offers great flexibility in modifying and incorporating targeting motifs.

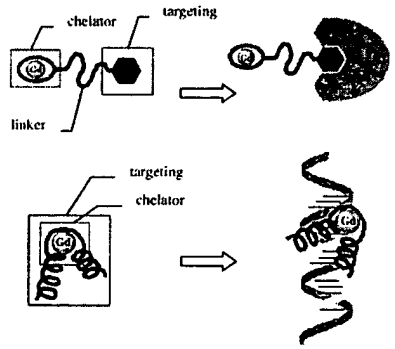

Fig. 1 Strategies for bifunctional contrast agent design. (top) Typical bifunctional contrast agent, with Gd-chelate and targeting moiety tethered by a linker. (bottom) Chimeric design, with Gd-chelate embedded within the targeting moiety.

H$_2$N-TERRRQQL**DKDGDGTIDERE**IKIWFQNKRAKIK-COOH

Scheme 1 Sequence of P3W, with Gd-binding EF-hand site shaded.

A helix–turn–helix (HTH) chimera, a metallopeptide based on the incorporation of an EF-hand Ca-binding loop into the DNA-binding HTH motif,[10,11] was investigated as a putative DNA sensing contrast agent. The peptide chimera P3W (Scheme 1) is based on the structurally similar α–α corner motifs of two unrelated proteins: *engrailed* homeodomain, a DNA-binding transcription factor, and *calmodulin*, a Ca-binding signaling protein.[10] Previous studies have shown that P3W folds upon binding one equivalent of trivalent lanthanide ions to the EF-hand Ca-binding site,[11] binds DNA,[12] and enters cells in the metallated form (unpublished data).[13] Circular dichroism and NMR studies of various Ln(III)P3W metallopeptides (Eu, Yb, La) determined that the folded structure of the metallopeptide is analogous to the parental α–α corner motifs.[11] The CD spectrum of GdP3W indicates the Gd-peptide is similarly well-folded in solution, and the affinity for Gd(III) is comparable to that of Eu(III) (log K_{EuP} = 5.21 ± 0.3; log K_{GdP} = 5.19 ± 0.4, pH 7.8). The EuP3W complex was found by Eu-luminescence spectroscopy to have two inner sphere water molecules ($q \approx 2.0$) both in the presence and absence of DNA,[14] suggesting that $q = 2$ for GdP3W as well.

The relaxivity of GdP3W was determined at 20 and 60 MHz (37 °C). The GdP3W metallopeptide was found to be a very efficient catalyst of water proton relaxation (Table 1). This relaxivity is significantly higher (6-fold at 60 MHz) than commercial agents, *e.g.* GdDTPA. In the presence of one equivalent of DNA, the relaxivity at 60 MHz increases by 100% to 42.4 mM^{-1} s^{-1}. The increase in r_1 upon DNA binding was expected, but the magnitude and unusual field dependence (Fig. 2) of the relaxivity were surprising. Unlike most slow tumbling complexes reported, the relaxivity at 60 MHz is significantly greater than at 20 MHz, notable because most clinical MRI scanners operate at 1.5 tesla (~65 MHz). It is clear from Fig. 2 that the relaxivity of the Gd–peptide–DNA complex is peaking near 60 MHz. This is likely a consequence of a difference in electronic relaxation imparted by the all oxygen donor set,[15] and

Table 1 Relaxivities (mM^{-1} s^{-1}) of GdP3W and GdDTPA (37 °C, 50 mM HEPES, pH 7.4) at 20 and 60 MHz in the presence and absence of one equivalent DNA (14-mer d.s. DNA: self-complementary 5′-GAGCTAATTAGCTC-3′)

	20 MHz		60 MHz	
Contrast agent	r_1	r_1 + DNA	r_1	r_1 + DNA
GdP3W	16.2	29.6	21.2	42.4
GdDTPA	4.6	4.9	3.5	3.5

† Electronic supplementary information (ESI) available: circular dichroism spectrum of GdP3W; additional phantom MR images; Experimental details. See http://www.rsc.org/suppdata/cc/b3/b307817e/

DOI: 10.1039/b307817e

suggests that this ligand environment may be a useful design feature in future contrast agents.[16,17] The relaxivity at 60 MHz is one of the higher relaxivities reported at this frequency, including other RIME type agents.[7,8] For instance, it is approximately double the relaxivity of the albumin targeted contrast agent MS-325 (23 mM^{-1} s^{-1} at 60 MHz).[7]

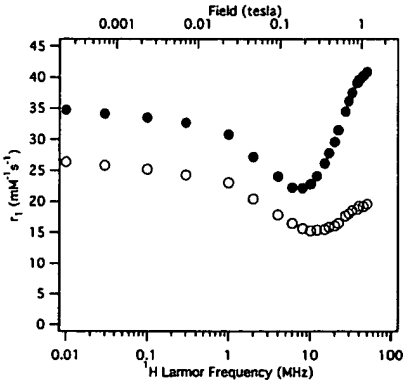

Fig. 2 ^{1}H NMRD relaxivities (35 °C) of GdP3W (33 μM) in the presence (filled) and absence (open circles) of 1 equivalent of DNA.

High relaxivity is a critical feature in the development of new contrast agents. Since MRI contrast agents are detected indirectly by their influence on water relaxation rates, relatively high concentrations of contrast agent are required. The benefit of a high relaxivity agent such as GdP3W is qualitatively shown in Fig. 3. T_1-weighted imaging (short T_1 giving positive contrast) at 1.5 T shows that at a fixed gadolinium concentration of 48 μM, the contrast between GdDTPA (**b**) and buffer (**a**) is difficult to discern. However, GdP3W is significantly brighter than the clinical agent. The DNA "switch" concept is shown by the addition of 1 equivalent of DNA to the GdP3W sample, resulting in a signal increase relative to GdP3W upon target binding (**d** *vs.* **c**). This emphasizes the importance of high relaxivity: if the target concentration is low, then it will be undetectable unless the relaxivity is high. An additional study (Supplemental) showed that the differences in relaxivity were borne out in the images: an 8.6 μM GdP3W + DNA phantom was isointense with 115 μM GdDTPA demonstrating that the same contrast can be achieved with 12 times less gadolinium. Control studies (Supplemental) demonstrated that addition of DNA did not enhance the signal of GdDTPA, and that DNA alone or P3W alone gave no signal enhancement. The electrostatic binding interaction between DNA and GdP3W could be salted out by a large excess of NaCl (> 150 mM) which caused a decrease in r_1 back to that of GdP3W alone. This was demonstrated in phantoms (Supplemental) or by measuring T_1 as a function of NaCl added. This demonstrates that the increase in GdP3W signal in the presence of DNA is truly due to reversible DNA-binding by the metallopeptide.[18]

This study shows that a $Gd(O)_8$ site in the context of a peptide-based DNA-targeting moiety can enhance T_1 contrast dramatically compared to commercial agents, and this contrast can be modulated by DNA-binding. Both the larger size (MWt > 4000 Da) causing a slower tumbling rate, and the two inner-sphere water molecules contribute to the increased relaxivity of GdP3W relative to GdDTPA. The dramatic increase in relaxivity upon target binding also suggests that incorporating rather than appending the Gd site to the targeting moiety is an effective approach to reducing rotational flexibility at the metal site. Whether this unique field dependence with increased sensitivity at higher fields is a generality for EF-hand-based metallopeptides, perhaps as a consequence of the unusual (for contrast agents) $Gd(O)_8$ coordination set, is under investigation. Although it is clear that the Gd affinity of the model HTH peptide P3W is too low for biological MRI applications, these experiments demonstrate a new class of bifunctional targeted MRI agents with an embedded Gd site, activated by and reporting on DNA-binding.

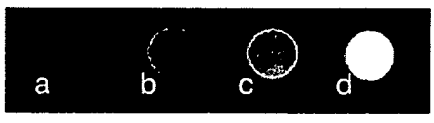

Fig. 3 T_1-weighted MR images of phantoms at 1.5 tesla in HEPES buffer, pH 7.4, ambient temperature. 3D SPGR TR = 40, TE = 4.0, α = 40°. *Images:* a) HEPES; b) 48 μM GdDTPA; c) 48 μM GdP3W; d) 48 μM GdP3W + DNA.

Notes and references

‡ Ed Parsons and Andrea Wiethoff are thanked for their help with the imaging studies, and Marga Spiller for recording the NMRD profiles. We thank the Center for Biocatalysis and Bioprocessing for NIH Predoctoral Training Grant T32 GM-08365 (JTW) and the National Science Foundation for a CAREER Award (CHE-0093000; SJF).

1 R. Weissleder and U. Mahmood, *Radiology*, 2001, **219**, 316.
2 P. Wunderbaldinger, L. Josephson and R. Weissleder, *Acad. Radiol.*, 2002, **9**, S304.
3 S. Flacke, S. Fischer, M. J. Scott, R. J. Fuhrhop, J. S. Allen, M. McLean, P. Winter, G. A. Sicard, P. J. Gaffney, S. A. Wickline and G. M. Lanza, *Circulation*, 2001, **104**, 1280.
4 S. Aime, L. Frullano and S. Geninatti Crich, *Angew. Chem., Int. Ed.*, 2002, **41**, 1017.
5 R. A. Moats, S. E. Fraser and T. J. Meade, *Angew. Chem., Int. Ed. Engl.*, 1997, **36**, 726; A. L. Nivorozhkin, A. F. Kolodziej, P. Caravan, M. T. Greenfield, R. B. Lauffer and T. J. McMurry, *Angew. Chem., Int. Ed.*, 2001, **40**, 2903; A. Y. Louie, M. M. Hüber, E. T. Ahrens, U. Rothbächer, R. Moats, R. E. Jacobs, S. E. Fraser and T. J. Meade, *Nat. Biotechnol.*, 2000, **18**, 321.
6 R. B. Lauffer, *Magn. Reson. Med.*, 1991, **22**, 339.
7 P. Caravan, N. J. Cloutier, M. T. Greenfield, S. A. McDermid, S. U. Dunham, J. W. M. Bulte, J. C. Amedio, Jr., R. J. Looby, R. M. Supkowski, W. D. Horrocks, Jr., T. J. McMurry and R. B. Lauffer, *J. Am. Chem. Soc.*, 2002, **124**, 3152.
8 L. M. De León-Rodríguez, A. Ortiz, A. L. Weiner, S. Zhang, Z. Kovacs, T. Kodadek and A. D. Sherry, *J. Am. Chem. Soc.*, 2002, **124**, 3514.
9 R. B. Lauffer, *Chem. Rev.*, 1987, **87**, 901.
10 Y. Kim, J. T. Welch, K. M. Lindstrom and S. J. Franklin, *J. Biol. Inorg. Chem.*, 2001, **6**, 173; J. T. Welch, M. Sirish, K. M. Lindstrom and S. J. Franklin, *Inorg. Chem.*, 2001, **40**, 1982.
11 J. T. Welch, W. R. Kearney and S. J. Franklin, *Proc. Natl. Acad. Sci. USA*, 2003, **100**, 3725.
12 Non-specific DNA-binding affinity: $K_d \leq 15$ μM; R. T. Kovacic, J. T. Welch and S. J. Franklin, *J. Am. Chem. Soc.*, 2003, **125**, 6656.
13 Uptake is presumably due to the membrane transduction domain sequence of engrailed. See: A. Prochiantz, *Curr. Opin. Cell Biol.*, 2000, **12**, 400.
14 Lifetimes (in ms) of the $^7F_0$–$^5D_0$ transition for EuP3W: τ_{H_2O} 0.35; τ_{D_2O} 2.11; EuP3W + DNA: τ_{H_2O} 0.41; τ_{D_2O} 2.49. S. Jain, J. T. Welch, W. D. Horrocks, Jr. and S. J. Franklin, *Inorg. Chem.*, 2003, in press.
15 Detailed analysis of the NMRD curves as a function of temperature, necessary to extract the molecular parameters defining relaxivity, is ongoing and will be reported later.
16 Excepting the hydroxypyridonate complexes developed by Raymond and coworkers (O_8 donor set), Gd complexes proposed as contrast agents typically have mixed N,O donor sets, and are based on polyaminocarboxylato ligands. The r_1 of GdP3W is much greater than Gd-TREN-1-Me-3,2-HOPO (also q = 2) and related complexes (~6–11 mM^{-1} s^{-1} at 60 MHz, 37 °C), presumably due to the embedded design, which limits the internal motion of the complex. Ref. 17.
17 J. Xu, S. J. Franklin, D. W. Whisenhunt, Jr. and K. N. Raymond, *J. Am. Chem. Soc.*, 1995, **117**, 7245; C. J. Sunderland, M. Botta, S. Aime and K. N. Raymond, *Inorg. Chem.*, 2001, **40**, 6746; S. Hajela, M. Botta, S. Giraudo, J. Xu, K. N. Raymond and S. Aime, *J. Am. Chem. Soc.*, 2000, **122**, 11228; S. M. Cohen, J. Xu, E. Radkov, K. N. Raymond, M. Botta, A. Barge and S. Aime, *Inorg. Chem.*, 2000, **39**, 5747; P. Caravan, J. J. Ellison, T. J. McMurry and R. B. Lauffer, *Chem. Rev.*, 1999, **99**, 2293.
18 The relaxivity of the GdP3W–DNA ternary complex (60 MHz, 37 °C) decreases as a function of NaCl, returning to the DNA-free value at > 150 mM added salt. Na^+ does not displace Ln^{3+} ions from P3W.

J|A|C|S
ARTICLES
Published on Web 05/07/2003

Sequence-Selective DNA Cleavage by a Chimeric Metallopeptide

Roger T. Kovacic, Joel T. Welch, and Sonya J. Franklin*

Contribution from the Department of Chemistry, University of Iowa, Iowa City, Iowa 52242

Received August 19, 2002; E-mail: sonya-franklin@uiowa.edu

Abstract: A chimeric metallopeptide derived from the sequences of two structurally superimposable motifs was designed as an artificial nuclease. Both DNA recognition and nuclease activity have been incorporated into a small peptide sequence. P3W, a 33-mer peptide comprising helices α2 and α3 from the engrailed homeodomain and the consensus EF-hand Ca-binding loop binds one equivalent of lanthanides or calcium and folds upon metal binding. The conditional formation constants (in the presence of 50 mM Tris) of P3W for Eu(III) ($K_a = (2.1 \pm 0.1) \times 10^5$ M^{-1}) and Ce(IV) ($K_a = (2.6 \pm 0.1) \times 10^5$ M^{-1}) are typical of isolated EF-hand peptides. Circular dichroism studies show that 1:1 CeP3W is 26% α-helical and EuP3W is up to 40% α-helical in the presence of excess metal. The predicted helicity of the folded peptide based on helix length and end effects is about 50%, showing the metallopeptides are significantly folded. EuP3W has considerably more secondary structure than our previously reported chimeras (Welch, J. T.; Sirish, M.; Lindstrom, K. M.; Franklin, S. J. *Inorg. Chem.* **2001**, *40*, 1982−1984). Eu(III)P3W and Ce(IV)P3W nick supercoiled DNA at pH 6.9, although EuP3W is more active at pH 8. CeP3W cleaves linearized, duplex DNA as well as supercoiled plasmid. The cleavage of a 5′-^{32}P-labeled 121-mer DNA fragment was followed by polyacrylamide gel electrophoresis. The cleavage products are 3′-OPO_3 termini exclusively, suggesting a regioselective or multistep mechanism. In contrast, uncomplexed Ce(IV) and Eu(III) ions produce both 3′-OPO_3 and 3′-OH, and no evidence of 4′-oxidative cleavage termini with either metal. The complementary 3′-^{32}P-labeled oligonucleotide experiment also showed both 5′-OPO_3 and 5′-OH termini were produced by the free ions, whereas CeP3W produces only 5′-OPO_3 termini. In addition to apparent regioselectivity, the metallopeptides cut DNA with modest sequence discrimination, which suggests that the HTH motif binds DNA as a folded domain and thus cleaves selected sequences. The de novo artificial nuclease LnP3W represents the first small, underivatized peptide that is both active as a nuclease and sequence selective.

Introduction

After decades of study, an understanding is emerging of cancer as a disease with its basis in regulatory gene mutations that prevent the correct self-regulation of growth.[1,2] The effect of these mutations might be mitigated if target gene sequences could be selectively cleaved or down-regulated. Recent advances in message profiling are leading to the possibility of identifying a small collection of overexpressed genes in tumerous tissues as targets for therapeutic down-regulation.[3] However, current small-molecule chemotherapy agents exhibit only subtle sequence preference based on shape complementarity and charge. These anticancer drugs are highly effective in certain contexts because of their ability to damage cancer cell DNA through covalent modification or double-stranded cleavage. These changes are difficult for the cell to repair, but there is little specificity in the sites targeted. As a consequence, there is substantial damage to nonmalignant cells as well. It is thus of great interest to develop selective chemotherapy agents that can target sequences of choice.

To design a sequence-specific DNA cleavage agent, both DNA-recognition and -reactive elements must be integrated. Transcription factors and repressor proteins are capable of selectively recognizing given duplex DNA sequences but do not cut DNA.[4] In contrast, hydrolytic cleavage of DNA by Lewis acid metal ions and complexes is well-known, but occurs with limited sequence discrimination. Although DNA is a robust molecule that is quite resistant to hydrolysis at neutral pH, lanthanide ions have been shown to be particularly efficient Lewis acids for promoting DNA cleavage.[5−7] We recognized that the supersecondary structure of the helix−turn−helix (HTH), a DNA-binding motif, and the EF-hand, a calcium-binding motif, are superimposable, (Figure 1) and exploited this similarity to design small peptides that bind lanthanides, retain the HTH fold, and cleave DNA.[8−10] These peptides represented the first examples of catalytically active EF-hands. Here we

(1) Abbot, A. *Nature* **2002**, *416*, 470−474.
(2) Hanahan, D.; Weinberg, R. A. *Cell* **2000**, *100*, 57−70.
(3) Ramsay, G. *Nat. Biotechnol.* **1998**, *16*, 40−44.
(4) Patikoglou, G.; Burley, S. K. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 289−325.
(5) Franklin, S. J. *Curr. Opin. Chem. Biol.* **2001**, *5*, 201−208.
(6) Hegg, E. L.; Burstyn, J. N. *Coord. Chem. Rev.* **1998**, *173*, 133−165.
(7) Sreedhara, A.; Cowan, J. A. *J. Biol. Inorg. Chem.* **2001**, *6*, 337−347.
(8) Kim, Y.; Welch, J. T.; Lindstrom, K. M.; Franklin, S. J. *J. Biol. Inorg. Chem.* **2001**, *6*, 173−181.
(9) Welch, J. T.; Kearney, W. R.; Franklin, S. J. *Proc. Nat. Acad. Sci., U.S.A.* **2003**, *100*, 3725−3730.

10.1021/ja0210998 CCC: $25.00

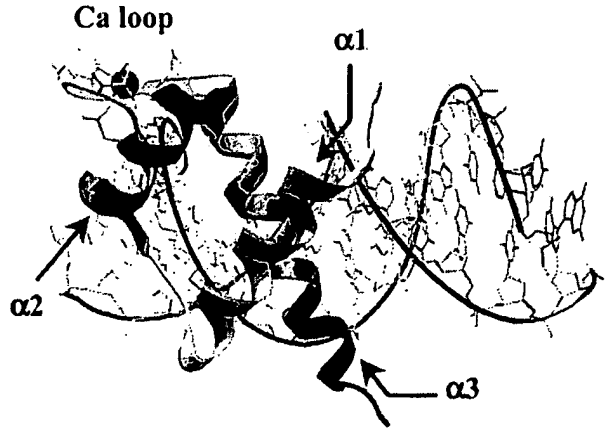

Figure 1. Computer-generated model of the interaction of chimeric peptides with DNA based on the HTH and EF-hand motifs from engrailed (red) and calmodulin (yellow), respectively. Crystal structures of engrailed homeodomain, cocrystallized with DNA (2HDD; gray and red ribbon, with blue DNA)[23] and one EF-hand of calmodulin (1OSA; yellow ribbon, with coordinated Ca(II)),[22] are overlaid to show the similarity of the folds. The peptides comprise α2 and α3 of engrailed and the Ca-binding loop of calmodulin.

report that this cleavage both results in regioselective products and is sequence dependent. Like the 3- and 4-α-helix bundle structures that are the foundation for a diversity of engineered metalloproteins,[11-13] the HTH motif retains a defined and predictable structure. Therefore the HTH is a robust structural scaffold for de novo protein design and the development of minimalist enzyme models. These de novo-designed HTH/EF-hand chimeric peptides are an initial step in the development of targeted DNA-binding metalloproteins as an alternative to antisense therapy or as a tool for manipulation of gene expression.

Materials and Methods

Materials. Chimeric peptides were designed on the basis of alignments of known crystal structures, using the freeware program SwissPDBViewer.[14] The design of the peptide P3 has been discussed in detail elsewhere.[8] Peptides (>95% pure) were obtained from New England Peptide (P3W) or the Caltech Peptide Synthesis Facility (P3). Sample peptide concentrations were based on extinction coefficients calculated from concentrations determined by peptide digestion (University of Iowa Peptide Facility; P3 $\epsilon_{280} = 254\ M^{-1}\ cm^{-1}$; P3W $\epsilon_{280} = 7289\ M^{-1}\ cm^{-1}$). $EuCl_3 \cdot 6H_2O$ and $(NH_4)_2Ce(NO_3)_6$ (anhydrous) were obtained from Sigma-Aldrich, and fresh stock solutions were made by weight before each use. Enzymes and buffers were obtained from New England Biolabs, and all other reagents were obtained from Sigma-Aldrich and used without further purification. All solutions were made with deionized distilled water (MilliQ 18 mΩ).

Plasmid DNA Cleavage. Plasmid pUC19 (New England Biolabs) was exchanged to Tris buffer containing no EDTA by affinity chromatography (Qiagen Mini-prep DNA spin column kit). EuP3W, CaP3W, and CeP3W (1:1) solutions were freshly prepared before use in Tris buffer. Samples containing 1.0 μg of plasmid, 50 mM Tris, pH 6.9 or pH 8.0 (at 37 °C) were incubated at 37 °C for 24 h with Eu(III) or Ce(IV) salts, or 1:1 metallopeptides. Prior to loading on the gel, samples were treated with Na^+-loaded Amberlyst cation-exchange resin to remove cations and peptide.[15] Products were analyzed by agarose gel electrophoresis (1%, in 1 × Tris-borate running buffer), stained with ethidium bromide, visualized by UV-transillumination, and photographed with a Kodak system 120 digital camera. Quantified supercoiled band intensities were normalized for weaker ethidium bromide staining by multiplying by 1.22.[16]

Oligonucleotide DNA Cleavage. pBR322 plasmid was cleaved at *BspH*I sites, 5′-^{32}P-labeled with T4 kinase protocol (New England Biolabs) and then further digested with *Ban*I. The 121 bp *BspH*I–*Ban*I oligonucleotide fragment was gel-purified and quantified by scintillation counting of the Cherenkov radiation (≤1 μM). Other pBR322 fragments (not shown) were analyzed in a similar manner. *Eco*R1/*Hae*III digested and phenol-extracted calf thymus DNA was used to raise the DNA concentration to 50 μM basepairs. Maxam–Gilbert sequencing reactions were prepared as size- and sequence markers by standard protocols.[17,18] For 3′-experiments, the 3′-strand of pUC19 restricted with *BspH*I was radiolabeled by incubating Klenow Fragment (3′ → 5′ exo-) DNA polymerase (New England Biolabs) and [^{32}P]-α-dCTP for 15 min in polymerase buffer. The reaction was quenched by adding EDTA and heating at 75° for 20 min. Unincorporated dCTP was removed with a G-50 ProbeQuant (Amersham Biotech) spin column. Subsequent *Ban*I restriction digestion gave the same fragment described above (with the opposite strand labeled).

Samples containing various concentrations of $(NH_4)_2Ce(NO_3)_6$ and peptides (50 mM Tris buffer, adjusted to pH 6.9 at 37 °C) were incubated at 37 °C for 24 h in an insulated heat block prior to analysis by polyacrylamide gel electrophoresis (20% acrylamide, 1 × TBE buffer under denaturing conditions), and an equal amount of radioactivity (counts) was loaded for each sample lane. Gels were exposed to the phosphor screens at −4 °C for up to 2 days prior to imaging with a Molecular Dynamics Storm Phosphoimager system. DNase 1[19] and Fe(II)-MPE[20] reactions were prepared by published methods. 3′-Labeled DNA experiments were carried out similarly, incubating $(NH_4)_2Ce(NO_3)_6$ or $EuCl_3$, and P3W (25 mM Tris buffer, adjusted to pH 7.2 at 37 °C) at 37 °C for 24 h prior to electrophoresis.

Circular Dichroism. CD spectra were recorded on an Olis-DSM17 spectrophotometer at 25 °C under N_2 atmosphere. Helical content was calculated from $[\Theta]_{222}$ as described previously.[8,21] Spectra were collected from 350 to 200 nm, with 0.5 point/nm, in a 0.1 cm cell. Samples in 50 mM Tris buffer, pH 6.9 (25 °C) were 50 μM in peptide.

Results and Discussion

Metallopeptide Design. The similarity in supersecondary structure of two physiologically unrelated motifs was utilized to design both activity and recognition into a single peptide domain. The HTH and EF-hand motifs consist of two helices at approximate right angles to one another, stabilized by conserved hydrophobic interactions along the helical inner surfaces. This "α–α corner" topology is a common structural motif in a diversity of proteins,[11,22] including DNA-binding proteins (HTH) and signaling or regulatory proteins (EF-hands). The remarkable similarity of the shape of these unrelated folds allows the turn to be modularly substituted. The metal-binding loop of the EF-hand thus promotes the α–α corner turn within the HTH sequence.

(10) Welch, J. T.; Sirish, M.; Lindstrom, K. M.; Franklin, S. J. *Inorg. Chem.* **2001**, *40*, 1982–1984.
(11) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. *Annu. Rev. Biochem.* **1999**, *68*, 779–819.
(12) Hill, R. B.; Raleigh, D. P.; Lombardi, A.; DeGrado, W. F. *Acc. Chem. Res.* **2000**, *33*, 745–754.
(13) Kennedy, M. L.; Gibney, B. R. *Curr. Opin. Struct. Biol.* **2001**, *11*, 485–490.
(14) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714–2723.
(15) Hettich, R.; Schneider, H.-J. *J. Am. Chem. Soc.* **1997**, *119*, 5638–5647.
(16) Hertzberg, R. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 313–315.
(17) Maniatis, T.; Fritsch, E. F.; Sambrook, J. *Molecular Cloning. A Laboratory Manual*, 2nd ed.; Cold Spring Harbor Laboratory Press: New York, 1982; Vols. 1–3, pp 475–478.
(18) Maxam, A. M.; Gilbert, W. *Methods Enzymol.* **1980**, *65*, 499–560.
(19) Galas, D. J.; Schmitz, A. *Nucleic Acids Res.* **1978**, *5*, 3157–3170.
(20) Dervan, P. B. *Methods Enzymol.* **2001**, *340*, 450–466.
(21) Chen, Y.-H.; Yang, J. T.; Martinez, H. M. *Biochemistry* **1972**, *11*, 4120–4131.
(22) Efimov, A. V. *FEBS Lett.* **1996**, *391*, 167–170.

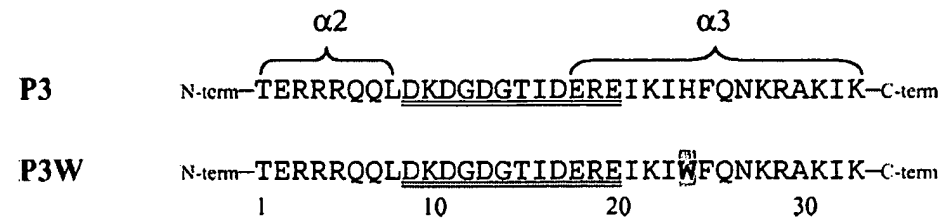

Figure 2. Sequences of P3W and P3, with helices α2 and α3 of engrailed indicated. The EF-hand Ca-binding loop is underlined, and the native Trp_{24} substituted in P3 with His_{24} is shaded.

The 33-mer peptides P3 and P3W were designed on the basis of overlaid crystal structures of the α−α corner motifs from calmodulin (1OSA)[23] and the engrailed homeodomain (1ENH).[24] The structures were oriented to align α−α corner helical axes of the two motifs (Figure 1, red and yellow motifs), from which the sequences of the chimeric peptides were derived. These peptides comprise helices α2 and α3 of engrailed, with the turn region replaced by the consensus Ca-binding EF-hand loop (Figure 2). Important hydrophobic interactions known to stabilize EF-hand structures[25] were retained at key positions before and after the loop (L_8, I_{21}, and W_{24} for peptide P3W), although the residues were derived from the engrailed sequence. The native α3 sequence of P3W preserves the hydrophobic core of the HTH motif and results in a well-structured metallopeptide, whereas P3, which has a Trp_{24} to His_{24} substitution, adopts a less defined tertiary structure.[9]

Solution Structure and Lanthanide Binding. The designed peptides P3 and P3W bind lanthanide ions via the consensus EF-hand site and fold to a nativelike structure. Both peptides P3 and P3W bind lanthanides with micromolar affinities under the sequencing gel electrophoresis conditions. The conditional binding affinity of P3W for Eu(III) and Ce(IV) was determined by fluorescence spectroscopy at pH 6.9, 50 mM Tris, following the intensity of the Trp_{24} emission at 350 nm as described previously.[26] A nonlinear least-squares fit of the intensity data for each metallopeptide supports a 1:1 Ln:peptide associative model. The conditional formation constants for Ce(IV) and Eu(III) with P3W were found to be $K_a = (2.6 \pm 0.1) \times 10^5$ M^{-1} and $K_a = (2.1 \pm 0.1) \times 10^5$ M^{-1}, respectively, which is very similar to the affinity of P3 for Eu(III) ($K_a = (1.2 \pm 0.5) \times 10^5$ M^{-1}) reported previously.[10,27] These values have been corroborated by luminescence titration experiments for EuP3 and EuP3W (W. D. Horrocks, Jr.; S. J. Franklin, unpublished data). As is typical for EF-hand sites, the affinity of P3W for Ca(II) ($K_a = (2.8 \pm 0.7) \times 10^4$ M^{-1}) is less than for the higher valent lanthanide ions.[25,28]

The chimeric metallopeptides exhibit solution behavior characteristic of folded domains. Previously we showed that P3 has little structure in the absence of lanthanides, but increased secondary structure upon metal binding.[8] In contrast, circular dichroism spectroscopy shows that P3W has a modest amount of secondary structure even as the apopeptide and folds further to a flexible, but more organized tertiary structure upon metal binding (Figure 3). The favorable impact of Trp_{24} (P3W) vs His_{24} (P3) on peptide stability and structure is noteworthy, and highlights the importance of hydrophobic side chains along the inner surfaces of the α−α corner motif in generating a structured domain. The shape and magnitude of the circular dichroism signals are typical of the increasing secondary structure changes upon metal binding observed for peptides derived from EF-hands.[29,30] Additionally, the CD spectrum of wild-type engrailed has the same relative CD band intensities, with a minimum at 208 nm and a weaker shoulder at 225 nm.[31] Thus, the CD signals of the metallopeptides are consistent with the spectra of the α−α corner parental motifs, and LnP3W is more folded than LnP3.

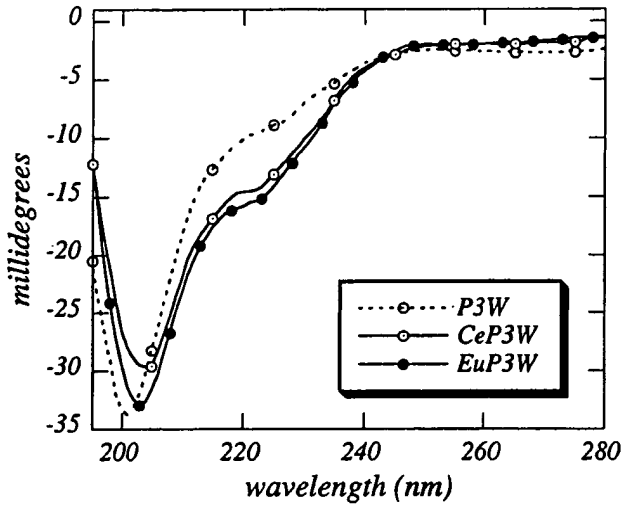

Figure 3. Circular dichroism spectra of peptide P3W (50 μM) in the presence of 3 equivalents of $EuCl_3$ or $(NH_4)_2Ce(NO_3)_6$. Spectra were collected on an Olis/DSM17 Spectrophotometer, at 25 °C and pH 6.9, in 50 mM Tris buffer, in a 0.1-mm cell. Under these conditions, estimates of α-helicity based on molar ellipticity $[\Theta]_{222}$ are 18% (P3W), 28% (CeP3W), and 30% (EuP3W).

An estimate of α-helical secondary structure based on molar ellipticity $[\Theta]_{222}$ from CD data[21] indicates that at pH 6.9 the CeP3W metallopeptide is 28% α-helical, and EuP3W is 30% helical (and up to 40% helical in the metal saturated form). These calculated estimates of secondary structure assume that all observed ellipticity at 222 nm is due to α-helical structure and that the fully folded species is 100% helical. Because this is not the case for either the EF-hand motif or HTH motif on which the peptide is based, the maximum amount of α-helicity predicted for a rigid fold would be 73% (24 of 33 residues occur in a putative helix). Further corrections based on the impact of helical length and end effects result in a predicted helicity of

(23) Chattopadhyaya, R.; Meador, W. E.; Means, A. R.; Quiocho, F. A. *J. Mol. Biol.* **1992**, *228*, 1177−1192.
(24) Kissinger, C. R.; Lui, B.; Martin-Blanco, E.; Kornberg, T. B.; Pabo, C. O. *Cell* **1990**, *63*, 579−590.
(25) Falke, J. J.; Drake, S. K.; Hazard, A. L.; Peersen, O. B. *Quart. Rev. Biophys.* **1994**, *27*, 219−290.
(26) Sirish, M.; Franklin, S. J. *J. Inorg. Biochem.* **2002**, *91*, 253−258.
(27) The affinity of P3 was determined by isothermal titration microcalorimetry, as it contains no tryptophan fluorophores.
(28) Drake, S. K.; Lee, K. L.; Falke, J. J. *Biochemistry* **1996**, *35*, 6697−6705.
(29) Procyshyn, R. M.; Reid, R. E. *J. Biol. Chem.* **1994**, *269*, 1641−1647.
(30) Reid, R. E.; Gariépy, J.; Saund, A. K.; Hodges, R. S. *J. Biol. Chem.* **1981**, *256*, 2742−2751.
(31) Ades, S. E.; Sauer, R. T. *Biochemistry* **1994**, *33*, 9187−9194.

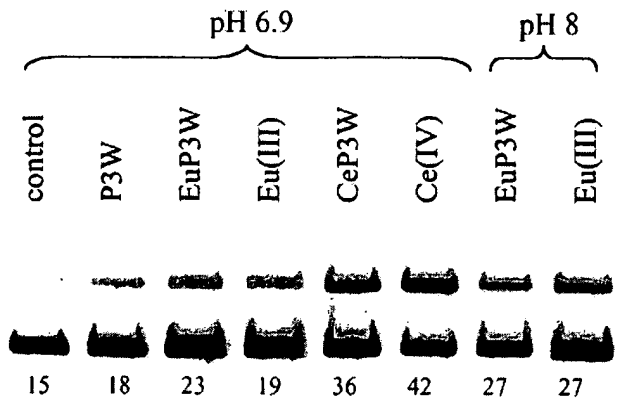

Figure 4. Agarose gel electrophoresis of supercoiled pUC19 plasmid DNA treated with peptide P3W and europium or cerium (50 mM Tris, pH as indicated). Lane 1: pUC19 plasmid control. Lane 2: 10 μM P3W (free). Lane 3: 10 μM EuP3W (1:1). Lane 4: 10 μM $EuCl_3$. Lane 5: 10 μM CeP3W (1:1). Lane 6: 10 μM $(NH_4)_2Ce(NO_3)_6$. Lane 7: 10 μM EuP3W (1:1). Lane 8: 10 μM $EuCl_3$. The open circular (upper) band is quantified below as a percentage of total DNA for this representative gel (±4% between repeated gels). The control lane at pH 8.0 (not shown) has the same background damage within ±1% of shown.

approximately 50% for a (time-averaged) fully folded metallopeptide.[32,33] As a comparison, the folded, wild-type engrailed homeodomain is reported to be 60% helical by CD spectroscopy.[31] Notably, an NMR solution structure model based on chemical shift index and NOE constraints has shown that the LaP3W complex retains the helix−loop−helix structure of both parent motifs.[9] These data suggest that the P3W metallopeptides are significantly folded and the chimeras can interact with DNA as a domain with tertiary structure.

DNA Cleavage Produces Phosphate Termini. The Eu(III) and Ce(IV) complexes of P3W catalyze the cleavage of supercoiled DNA, producing primarily single cuts (open circular product) and demonstrating that an exposed EF-hand is catalytically competent. As observed for EuP3,[10] in repeated gels EuP3W nicks supercoiled DNA slightly more efficiently than the uncomplexed Eu(III) ions (Figure 4, lanes 3 and 4, respectively), although EuP3W is more active at pH near 8 (lane 7). Both CeP3W and free Ce(IV) ions (lanes 5 and 6) are more active than the Eu(III) catalysts, even near neutral pH. This is consistent with the Lewis acidity, and therefore nucleophilic activity predicted based on the pK_a of inner sphere water molecules for each complex.[5,34,35]

The greater reactivity of Ce(IV) chimeras toward supercoiled DNA prompted us to focus on these complexes initially for sequencing reactions. ^{32}P-labeled DNA oligonucleotides (83−153 bp fragments) were incubated with $(NH_4)_2Ce(NO_3)_6$ and varying concentrations of peptides P3 or P3W at 37 °C for 24 h and analyzed by polyacrylamide gel electrophoresis. The strain inherent in supercoiled plasmid is not required to achieve cleavage, as we observe cutting of linearized ^{32}P-labeled DNA oligonucleotides as well. Although the Eu(III) experiments gave less product, the hydrolytic mechanism of cleavage is well established[6,7,36,37] and gives a nice reference to compare the more active Ce(IV) catalysts. The Eu(III) and Ce(IV) metallopeptides have essentially identical product profiles in the gels, although Eu(III) cleavage is much weaker. We find the cleavage gives single products (phosphates), likely occurring via a hydrolytic rather than an oxidative mechanism, and that peptide−DNA binding protects or suppresses cleavage by free metal.

The hydrolytic mechanism of DNA cleavage by Ce(IV) salts has recently been demonstrated by several groups.[35,38−41] Komiyama and co-workers showed that aqueous Ce(IV) salts hydrolytically cleave deoxyribodinucleotides, by analyzing the products by HPLC. They further demonstrated that DNA oligonucleotides were also hydrolytically cleaved with no dependence on molecular oxygen, generating termini that could be enzymatically manipulated. The cleavage was determined to be catalyzed by tetravalent, and not trivalent cerium. This analysis was elaborated by Branum and Que,[38,40] who demonstrated that ligand chelates of Ce(IV), as well as aqueous Ce(IV) salts, cleave DNA via a hydrolytic rather than oxidative mechanism, generating only hydroxyl and phosphate termini.

Hydrolysis by free Ce(IV) ion[42] produces both 3′- and 5′-phosphate termini, as there is no stereoselectivity in the hydrolysis mechanism. For the 5′-^{32}P-labeled fragment shown in Figure 5, two bands per base are generated by Ce(IV) ions (lane 6) under our experimental conditions. As observed by Que and by Komiyama, the resultant termini comigrate with the Maxam−Gilbert fragments (lanes 1−3; p = 3′-OPO_3) and DNase I fragments (lane 4; h = 3′-OH).[43] Both 3′-products are readily observed, although the bands due to 3′-OH termini are more intense. Oxidative cleavage by Fe(II) methidiumpropyl-EDTA (MPE) and peroxide generates both 3′-OPO_3 and 3′-phosphoglycolate termini (and corresponding 5′-products) (lane 5, g = 3′-phosphoglycolate).[44] One band (p), but not both, of the free Ce(IV) products comigrates with these 4′-oxidative products, which confirms the 3′-OPO_3 band assignment and the non-oxidative mechanism of uncomplexed Ce(IV) ion cleavage.

We have also investigated the 5′-termini products by radiolabeling the 3′-terminus of the same region (opposite strand; Supporting Information Figure S2). In this case, Maxam−Gilbert, DNaseI, and oxidative Fe-MPE/H_2O_2 cleavage all produce 5′-OPO_3 termini (p). Ce(IV) alone again leaves 5′-OPO_3 (p; major band) and 5′-OH (h; minor band), which complements both the observed termini and their intensities across the cut (Figure 5). Eu(III), which is well-known to cleave only by hydrolysis, gives the same OPO_3 and OH products at both 5′- and 3′-termini as observed for Ce(IV). This again corroborates the hydrolytic mechanism of uncomplexed Ce(IV) DNA cleavage.

In the presence of CeP3W and CeP3, only a single product is observed at each cleaved base step (Figure 5, lane 7). The products of chimera-catalyzed cleavage comigrate with the bands in Maxam−Gilbert sequencing lanes (5′-labeled DNA), which

(32) Hirst, J. D.; Brooks, C. L. I. *J. Mol. Biol.* **1994**, *242*, 173−178.
(33) The structure contains two separated helices. If two residues on either end of each helix do not contribute significantly to $[\Theta]_{222}$ as described in ref 32, the number of contributing residues is reduced to 16 of 33, or 48%.
(34) Kim, Y.; Franklin, S. J. *Inorg. Chim. Acta* **2002**, *341*, 107−112.
(35) Komiyama, M. *Chem. Commun.* **1999**, *16*, 1443−1451.
(36) Ott, R.; Krämer, R. *Appl. Microbiol. Biotechnol.* **1999**, *52*, 761−767.
(37) Schneider, H.-J.; Rammo, J.; Hettich, R. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1716−1719.
(38) Branum, M. E.; Tipton, A. K.; Zhu, S.; Que, L., Jr. *J. Am. Chem. Soc.* **2001**, *123*, 1898−1904.
(39) Komiyama, M.; Takeda, N.; Takahashi, Y.; Uchida, H.; Shiiba, T.; Kodama, T.; Yashiro, M. *J. Chem. Soc., Perkin Trans. 2* **1995**, 269−274.
(40) Branum, M. E.; Que, L., Jr. *J. Biol. Inorg. Chem.* **1999**, *4*, 593−600.
(41) Sumaoka, J.; Azuma, Y.; Komiyama, M. *Chem. Eur. J.* **1998**, *4*, 205−209.
(42) "Free metal ions" describe ions not complexed by peptide, although they exist as aquo/hydroxide Tris-stabilized species under these conditions.
(43) Only the former fragment is observed with 5′-^{32}P-labeled oligonucleotides.
(44) Burrows, C. J.; Muller, J. G. *Chem. Rev.* **1998**, *98*, 1109−1151.

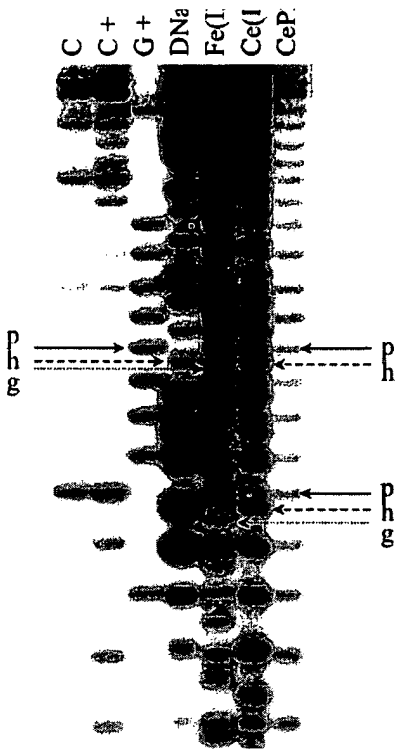

Figure 5. Polyacrylamide gel electrophoresis of cleavage products of a 5′-end-labled 121 bp DNA oligonucleotide fragment. Cleavage by Ce(IV) ions produces only the hydrolytic products 3′-OPO_3 and 3′-OH. The chimera CeP3W cleaves DNA, producing only 3′-OPO_3 termini. Lanes 1−3: Maxam−Gilbert sequencing lanes (p = 3′-OPO_3 termini). Lane 4: DNase I treated DNA (h = 3′-OH termini). Lane 5: Fe(II)MPE treated DNA (p = 3′-OPO_3 and g = 3′-phosphoglycolate termini). Lane 6: DNA incubated at 37 °C, pH 6.9 for 24 h with 30 μM $(NH_4)_2Ce(NO_3)_6$. Lane 7: DNA incubated at 37 °C, pH 6.9 for 24 h with 30 μM CeP3W.

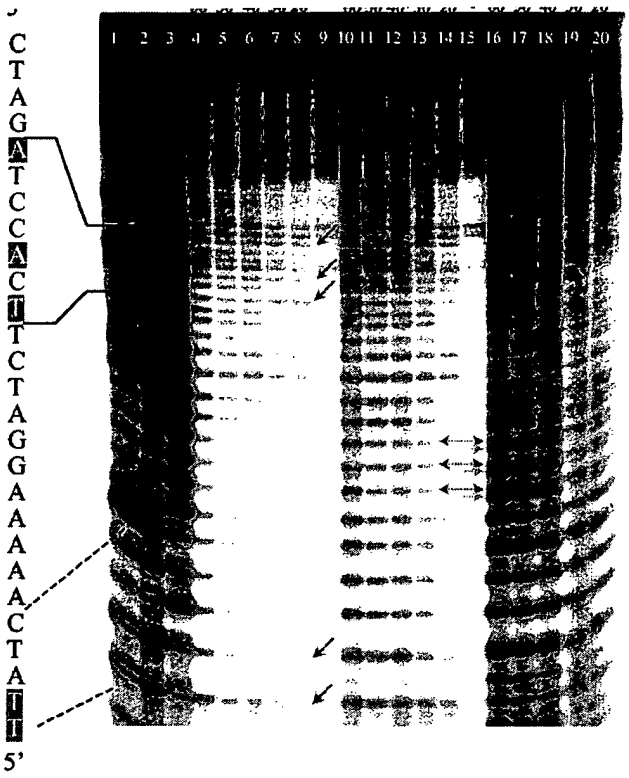

Figure 6. Polyacrylamide gel electrophoresis: the effect of metal−peptide concentrations on DNA cleavage. ^{32}P-Labeled oligonucleotides were incubated with 40 μM of peptides P3 or P3W and increasing concentrations of $(NH_4)_2Ce(NO_3)_6$ (from 20 to 60 μM, as indicated) at 37 °C for 24 h, pH = 6.9, 50 mM Tris buffer. Samples with no peptide or metal served as controls, and Maxam−Gilbert sequencing reactions as sequence markers. Bands that are notably enhanced over controls are indicated with green arrows for P3W lanes. Horizontal red hashed and orange solid arrows indicated 3′-phosphate and 3′-hydroxyl termini, respectively. Both bands were observed in the metal-only lanes, but only 3′-phosphate products were observed in the peptide-containing lanes.

suggests the metallopeptides are generating 3′-OPO_3 termini exclusively. In contrast, Que's bimetallic polyaminocarboxylate nuclease generated exclusively 3′-OH termini.[45] However, the complementary cleavage product (3′-labeled DNA) shows that Ce(IV)P3W also produces only 5′-OPO_3 termini, instead of the expected 5′-OH, implying the loss of the intervening sugar moiety (by either hydrolysis or oxidative cleavage), rather than a simple single-strand cut. It is possible that the peptide has a relatively long residence time on the DNA, allowing a second *exo*-nuclease step, once the initial *endo*-nuclease cleavage has been accomplished. Investigations to determine the reason for the unusual phosphate termini produced under various conditions are ongoing.

Peptide binding apparently also protects the DNA from further random cleavage by excess free metal in analogy to footprinting studies, although at higher metal:peptide ratios (≥4:1), the double band cleavage pattern generated by uncomplexed Ce(IV) hydrolysis is again observed. These random interactions are likely enhanced by the low-salt conditions, chosen to maximize DNA−peptide interactions, and prevent competition by Ca(II) or Mg(II) for the peptide.[46]

Sequence-Selective DNA Cleavage. In addition to single OPO_3 products, the chimeric peptides direct sequence-dependent cleavage. As shown in Figure 6, CeP3W-catalyzed cutting is not sequence-neutral, but rather selected sequences are preferentially cleaved (arrows). This nonrandom cleavage pattern suggests that the chimera occupies a set of higher affinity sites for greater average residence times, resulting in more damage at those positions. Sequence discrimination is observed to a lesser extent for CeP3 (Figure 6, lanes 10−14). CeP3 lacks important internal hydrophobic contacts (from Trp_{24}) that stabilize the HTH structure and is thus a less organized domain than CeP3W.[9]

These cleavage patterns are consistent with preferential binding by a folded domain at selected sites, such as those indicated by the arrows and shaded sequences in Figure 6. On this particular fragment, cleavage at the 5′-thymidine of the 5′-TCACCT-3′ sequence (black arrow) is notably more intense.

(45) Because the products have 3′-OPO_3 termini, present in both oxidative and hydrolytic reactions, it cannot be ruled out that the cleavage mechanism by Ce(IV) in the presence of peptide is oxidative, although cleavage by the free Ce(IV) ion and by Eu(III) is not. Furthermore, the analysis of the EuP3W termini is complicated by the relatively weak cleavage compared to Ce(IV). Consequently, cleavage intensities are barely over background, and although the products of EuP3W hydrolysis appear to be OPO_3 in both directions (paralleling the Ce(IV) results), the data are inconclusive.

(46) No change in background occurs during the course of the reaction, as 0 h and 24 h control lanes look identical.

However, the length and sequence of the global consensus recognition site remains to be determined.

It is notable that a very different apparent sequence dependence by EuP3W can be imposed than the T/C-rich sequence targeted under neutral cleavage conditions. Incubating the reactants under slightly acidic conditions (< pH = 6.5) partially depurinates the DNA. The metallopeptides (both EuP3W and CeP3W) then target and preferentially hydrolyze these abasic A and G steps (again leaving OPO_3 termini; see Supporting Information Figure S3).

The P3W chimera represents a "half domain" relative to the parental engrailed DNA-binding homeodomain, as the N-terminal tail and first helix have been omitted (Figure 1, gray ribbon). Nonetheless, the peptide is capable of DNA-binding with sequence discrimination, although the recognition site is not necessarily expected to be that of the parent. Engrailed recognizes the 6-bp sequence 5′-TAATTA-3′ as a monomer, with strong binding affinity (less than 1 nM).[31] The consensus general homeodomain binding site is of the form 5′-TAATNNN-3′, comprising 6−7 base pairs, and in many cases encompasses a "core" 5′-TAAT-3′ sequence.[47] Homeodomains typically interact with DNA via base contacts to both the major groove (the recognition helix, α3) and the minor groove (N-terminal tail and α1), along with phosphate backbone contacts from both α2 and α3 (Figure 1). The recognition helix contacts the major groove within the 3′-region of the recognized sequence (5′-TAATNN-3′), and the N-terminal tail contacts the minor groove along the 5′ region of the sequence (5′-TAATNN-3′).[47] In the native engrailed homeodomain, α2 is on the "backside" of the domain, and does not make direct base contacts. In contrast, our EF-hand/HTH chimeric peptides do not include the N-terminal tail and α1 (Figure 1, gray helix), so specific interactions with the 5′-TAA sequence of the engrailed recognition site are less likely. However, the engrailed homeobox target sequence is an overlapping palindrome, so the 3′-region that the recognition helix contacts is 5′-ATTA-3′, the coincidental complement to the "core" sequence. Several fragments of pUC19 were cleaved and labeled near 5′-TAAT-3′ sequences (data not shown), but no preferential cleavage at these sites was observed.

Unsurprisingly, the selectivity of the peptides differs from engrailed; the peptide chimera is sequence *selective*, although not sequence *specific* like the native homeodomain. Without the anchoring tail, α3 may be rotated in major groove, altering the relative orientation of α2 and allowing contacts to bases on the 3′-side of the site that are not accessible to the full domain. However, the metallopeptide appears to be binding to DNA as a structured unit to discriminate among sequences. The influence of DNA flexibility, flanking sequences, and length of the recognition site remains to be understood.

The cleavage rates and efficiency that we observe are typical for molecular hydrolase models.[5−7] The rate of EuP3-catalyzed bis-nitrophenyl phosphate cleavage, a DNA substrate analogue, was found to be 0.1 M^{-1} s^{-1}, which represents a hydrolysis rate enhancement of approximately 10^6 over uncatalyzed cleavage.[10] Although the cleavage rates, and consequently, band intensities are lower than often observed for small-molecule oxidative or photocleavage nucleases,[48,49] variations in band intensities are readily observed at low CeP3W concentrations. These variations cannot be explained by nonspecific binding and cleavage alone. While hydrolysis rates, and thus cleavage products, are enhanced at higher metallopeptide concentrations (above 50 μM peptide), competitive, nonspecific DNA binding becomes a factor, masking specific binding and cleavage.[50] Under such conditions, nonspecific (but still apparently phosphate-directing) cleavage by the metallopeptides results in ladder-like product distributions.

The level of sequence-dependent cleavage promoted by the chimeras is unprecedented in an underivatized peptide of this size. The one earlier report of selective Ln-mediated hydrolysis involved antisense recognition, by a 19-mer DNA oligonucleotide with an appended 5′-iminodiacetate-chelated Ce(IV) ion.[51] Barton and co-workers have also reported examples of small, zinc-binding peptides (16−22 residues) that hydrolyze DNA with sequence discrimination,[52,53] although in that case, the peptide is augmented by an appended Rh-intercalator complex that delivers and orients the peptide within the DNA groove. Additionally there are several oxidatively active examples in which selective damage was conferred by larger protein domains. Isolated DNA-binding domains such as Fos and Sp1 incorporating an amino-terminal NH_2-Xaa-Xaa-His Cu(II) or Ni(II) chelating sequence have been found to bind DNA sequence-selectively, thus generating diffusible radicals adjacent to localized sequence targets.[54−56] Similarly, Sarkar and co-workers reported sequence-dependent oxidative DNA damage by an iron-loaded estrogen receptor DNA-binding domain (84-mer), which comprises two zinc finger motifs per monomer.[57]

Even without reactivity, small peptides that fold and interact with DNA are unusual. There are very few examples of small, well-folded "miniproteins," natural polypeptides 20−42 amino acids in length with defined tertiary structure,[58−63] and even fewer that bind DNA selectively. A notable example is Schepartz's 42-residue avian pancreatic peptide (aPP), which was modified to include leucine zipper basic region DNA contacts.[63] Several designed single zinc finger peptides have also been shown to have affinity for selected RNA sites.[59,64] While these peptides bind strongly and selectively to targets, they were not designed for reactivity. In an alternative approach, Berg and co-workers redesigned a zinc finger peptide (22

(47) Ledneva, R. K.; Alexeevskii, A. V.; Vasil'ev, S. A.; Spirin, S. A.; Karyagina, A. S. *Mol. Biol.* **2001**, *35*, 647−659.

(48) Dervan, P. B. *Science* **1986**, *232*, 464−471.
(49) Erkkila, K. E.; Odom, D. T.; Barton, J. K. *Chem. Rev.* **1999**, *99*, 2777−2795.
(50) The random binding affinity for LnP3W species for DNA has been estimated from plasmid gel shift assays to be ~15 μM, which is in good agreement with similar studies on EuP3 (ref 8).
(51) Komiyama, M. *J. Biochem.* **1995**, *118*, 665−670.
(52) Fitzsimons, M. P.; Barton, J. K. *J. Am. Chem. Soc.* **1997**, *119*, 3379−3380.
(53) Copeland, K. D.; Fitzsimons, M. P.; Houser, R. P.; Barton, J. K. *Biochemistry* **2002**, *41*, 343−356.
(54) Harford, C.; Narindrasorasak, S.; Sarkar, B. *Biochemistry* **1996**, *35*, 4271−4278.
(55) Long, E. C. *Acc. Chem. Res.* **1999**, *32*, 827−836.
(56) Nagoaka, M.; Hagihara, M.; Kuwahara, J.; Sugiura, Y. *J. Am. Chem. Soc.* **1994**, *116*, 4085−4086.
(57) Conte, D.; Narindrasorasak, S.; Sarkar, B. *J. Biol. Chem.* **1996**, *271*, 5125−5130.
(58) Dahiyat, B. I.; Mayo, S. L. *Science* **1997**, *278*, 82−87.
(59) Friesen, W. J.; Darby, M. K. *J. Biol. Chem.* **2001**, *276*, 1968−1973.
(60) Gellman, S. H.; Woolfson, D. N. *Nat. Struct. Biol.* **2002**, *9*, 408−410.
(61) Hill, R. B.; DeGrado, W. F. *J. Am. Chem. Soc.* **1998**, *120*, 1138−1145.
(62) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425−430.
(63) Zondlo, N. J.; Schepartz, A. *J. Am. Chem. Soc.* **1999**, *121*, 6938−6939.
(64) McColl, D. J.; Honchell, C. D.; Frankel, A. D. *Proc. Nat. Acad. Sci., U.S.A.* **1999**, *96*, 9521−9526.

residues) with an open coordination site, which could potentially have hydrolytic nuclease activity.[65] Berg found this minimalist zinc finger folded correctly, though it did not show activity for acetate or aldehyde hydrolysis, and a test of nuclease activity was not reported.

Our de novo designed lanthanide chimera represents the first underivatized small peptide that is a sequence-selective artificial nuclease. The chimeras illustrate how a robust fold can be redesigned to incorporate catalytic activity into a minimalist recognition motif. The recognition helix of the HTH motif delivers a Lewis acidic ion (Eu(III) or Ce(IV)) to the DNA backbone for cleavage with modest sequence discrimination. Whether our chimeric peptides have specific targets, for which they exhibit higher selectivity and affinity, is under investigation. The combination of folding, DNA binding, and sequence-dependent cleavage makes chimeras based on the HTH scaffold a potential foundation for de novo nuclease design, perhaps leading to new clinical and biochemical tools.

Acknowledgment. This work was supported by NSF CAREER Award, CHE-0093000, and a Center for Biocatalysis and Bioprocessing NIH Predoctoral Training Grant, T32 GM-08365 (J.T.W.).

Supporting Information Available: Polyacrylamide gel electrophoresis: DNA cleavage by 1:1 CeP3W (25−45 μM); 5′- and 3′-^{32}P-radiolabeled DNA cleavage by CeP3W and EuP3W; EuP3W cleavage products of DNA under acidic conditions (depurinated DNA) (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA0210998

(65) Merkle, D. L.; Schmidt, M. H.; Berg, J. M. *J. Am. Chem. Soc.* **1991**, *113*, 5450−5451.

ELSEVIER

Journal of Inorganic Biochemistry 91 (2002) 253–258

JOURNAL OF Inorganic Biochemistry

www.elsevier.com/locate/jinorgbio

Hydrolytically active Eu(III) and Ce(IV) EF-hand peptides

Mallena Sirish, Sonya J. Franklin*

Department of Chemistry, University of Iowa, Iowa City, IA 52242, USA

Received 15 November 2001; received in revised form 5 February 2002; accepted 16 February 2002

Abstract

A chimeric peptide (P4) has been designed to incorporate an EF-hand metal-binding loop into the context of the helix-turn-helix DNA binding motif of the engrailed homeodomain. This construct binds lanthanides, and in the presence of these metals, promotes the cleavage of supercoiled DNA and model phosphate esters (bisnitrophenyl phosphate). P4 binds lanthanides with moderate affinities (Eu(III), log K_a=4.85; and Ce(IV), log K_a=5.23). The structure of P4 is enhanced by metal binding, but the increase in secondary structure observed by CD is small, and suggests the metallopeptide is also quite flexible. Despite this flexibility, the efficient cleavage of DNA at low concentrations is dependent on the metallopeptide, and not on peptide or metal alone. This enhanced reactivity suggests the designed DNA-binding EF-hand peptides deliver the metal to the DNA for catalysis, even without rigid secondary structure. © 2002 Published by Elsevier Science Inc.

Keywords: Lanthanides; Hydrolysis; De novo design; EF-hand; Nuclease

1. Introduction

There has recently been rapid progress in engineering metallopeptides as minimalistic active-site models for enzymes, building on scaffolds such as α-helical bundles [1,2]. Another promising direction for de novo metallopeptide design is incorporating an existing metal-binding site into a native protein fold, an approach that is gaining interest for protein redesign [3]. In this vein, we are designing functional metallopeptides based on two structurally similar protein folds, which both have the topology classified as an α–α corner motif [4]. This corner comprises two orthogonal helices, and is found in diverse protein contexts. Both the helix-turn-helix (HTH), a DNA binding motif, and the EF-hand, a Ca-binding motif, exhibit this α–α corner structure (Fig. 1). Additionally, the lanthanides are known to bind to EF-hand sequences with the same geometry as Ca(II) ions, and to be catalysts of hydrolytic DNA cleavage. There are several examples of small molecule lanthanide complexes that cleave DNA efficiently, showing the versatility of this lewis acid in DNA hydrolysis [5–9]. Therefore, the similarity in the HTH and EF-hand motifs' native structure can be utilized to design peptides that bind both metals and DNA, as nucleases with the potential for tunable specificity [10,11].

The sequence of the chimeric 33-mer peptide P4 is based on the overlay of two physiologically unrelated α–α corner motifs, derived from regions of the proteins calmodulin and engrailed. This peptide was found to have only modest structure upon metal binding. The aim of this

Fig. 1. Two α–α corner motifs, from which the P4 peptide sequence is derived. Top: third EF-hand of calmodulin (1OSA). Bottom: Helix-turn-helix motif of engrailed homeodomain (1ENH).

*Corresponding author. Tel.: +1-319-353-2244; fax: +1-319-335-1270.

E-mail address: sonya-franklin@uiowa.edu (S.J. Franklin).

0162-0134/02/$ – see front matter
PII: S0162-0134(02)00408-7

work was to explore the nuclease activity of a flexible metallopeptide, and to investigate this activity with Ce(IV) as well as the trivalent lanthanide Eu(III) ions. P4, like the previously reported chimeric peptides [10,11], binds trivalent lanthanide metals, and catalyzes the cleavage of DNA and bisnitrophenyl-phosphate esters. This peptide also binds Ce(IV), and is the first example of a catalytically active Ce(IV)-EF-hand motif. Both the Eu(III) and Ce(IV) P4 metallopeptides are most active toward plasmid cleavage at concentrations near K_d. The effect of the sequence on lanthanide metal affinity, peptide flexibility, and hydrolytic reactivity is an important factor in future designs.

2. Experimental

2.1. Peptides and reagents

The 33-residue peptide P4 was synthesized and purified by Anaspec, San Jose, CA (P4: NH_3-TERRRFRVFDKDGNGYISAAEKIWFQNKRAKIK-COOH). Concentrations were determined from the measured extinction coefficient (P4 ε=6484 M^{-1} cm^{-1} at 278 nm), calculated by peptide digestion and quantification (University of Iowa Peptide Synthesis Facility). Eu(III) ions were added as the chloride salts (Aldrich, 99.99%), and Ce(IV) added as $(NH_4)_2Ce(NO_3)_6$ (Aldrich, 99.99%), from 10-mM stock solutions.

2.2. Fluorescence titrations

The titration of P4 with various metals was monitored using fluorescence spectroscopy. The intrinsic changes in the tryptophan fluorescence (W_{24}) upon the addition of metals was monitored on an Aminco-Bowman Series 2 Fluorimeter, with an excitation wavelength of 290±0.5 nm. Titrations were carried out in 10 mM Tris buffer, pH 7.9±0.2 by adding concentrated stock solutions of $EuCl_3$ or $(NH_4)_2Ce(NO_3)_6$ (1 mM) containing 50 μM of P4, to an equally concentrated solution of P4 in a 1-cm cuvette (peptide concentration is invariant with volume). Titrations were repeated three times, and each carried out until no further spectral changes were detected. The drop in emission intensity as a function of metal was observed at several wavelengths from 340 to 360 nm. These data were iteratively fit to a single exponential curve using a linear non-least squares algorithm, and were well described with a 1:1 association model [12,13]. Additional equilibria did not improve the fit. It should be noted that this association constant reflects peptide competing for a Tris-stabilized aqua or hydroxide complex (Ln-Tris $\log K_a$=2.44) [14]. Error limits in $\log K$ are estimated at ±3% based on repeated measurements.

2.3. Circular dichroism spectroscopy

Circular dichroism (CD) spectra were recorded on a dual beam Aviv 60DS spectrophotometer at 25 °C. Peptide samples (50 μM) in 10 mM Tris buffer, pH 7.0, were scanned from 260 to 200 nm, with signal averaging (1.0-nm bandwidth, 0.5-nm resolution, 0.1-cm pathlength). The titrations of free and metallated peptide with trifluoroethanol (TFE) were followed under similar conditions. Aliquots of a stock TFE solution (95%) containing free P4 peptide or 1:1 Eu:P4 (50 μM) were added to the buffered samples, so the overall concentration of peptide and metal was invariant over the course of the titration.

The percent helicity of the peptide under various conditions was estimated from the molar ellipticity at 222 nm ($[\Theta_{222}]$), though this number assumes only α-helical structure contributes to the intensity at 222 nm [10,15].

2.4. DNA plasmid cleavage gels

The cleavage of supercoiled DNA was followed by agarose gel electrophoresis. Unless noted otherwise, samples containing 10–20 μM metal or 1:1 metal-peptide reagents in 50 mM Tris, pH 7.0, were incubated with pBR322 plasmid DNA for 24 h at 37 °C in a covered heating block to prevent solvent evaporation. The reaction mixtures were treated with Na^+-loaded Amberlyst cation exchange resin for 30 min at 37 °C, followed by centrifugation, in order to complex the highly positively charged peptides and allow the observation of DNA products in the gel [16]. Agarose gels (1% agarose in 0.5×TBE buffer) were stained with ethidium bromide overnight (0.5 μg/ml), and the gels were visualized by UV-transillumination and photographed with a Kodak System 120 Digital Camera (1.2 million pixels). Gel experiments were conducted within the linear range of the camera, which was determined from gels of partially digested DNA standards over a range of concentrations. Bands were quantified using Molecular Dynamics Image Quant Software. The quantified values for supercoiled DNA were multiplied by 1.22, to correct for the reduced staining of this structural form by ethidium bromide [17]. Numbers reported in Fig. 4 reflect averages from three repeated gels.

2.5. Bisnitrophenyl phosphate (BNPP) cleavage assays

The rate of bisnitrophenyl phosphate cleavage was followed by UV–visible spectroscopy as reported previously [11], observing the formation of nitrophenylate product at 400 nm. Data were collected on a dual-beam Varian Cary 300 spectrophotometer with Peltier temperature-controlled jacketed cell holder, using 1-cm cells. All measurements were made under turn-over conditions (500 μM BNPP) at 37 °C in 10 mM Tris buffer, pH 7.0,

from freshly prepared stock solutions. Samples contained $EuCl_3$ and P4 at various concentrations. Equimolar BNPP (500 μM) was also added to each reference cuvette to correct for any residual absorbance associated with the BNPP at 400 nm. The absorbance at 400 nm was recorded over the initial 12–15 h of reaction (≤10% BNPP converted), and plotted versus time to give the pseudo first order rate constants (Table 1). No measurable hydrolysis of BNPP was observed in the absence of metal or peptide.

3. Results and discussion

The peptide P4 was designed based on structural overlays of calmodulin [18] (PDB code 1OSA) and engrailed homeodomain [19] (PDB code 1ENH), in analogy to previously reported peptides P4a and P3 (Fig. 1; Scheme 1) [11]. Residues 6–21 of P4 are derived from the calmodulin sequence, which in the native structure comprises the Ca-loop and includes a full α-turn of both N- and C-terminal helices. In native Ca-binding EF-hand proteins [20–22] the affinity of the Ca-binding loop for Ca(II) and Ln(III) ions is dependent in part on the peripheral sequences bracketing the Ca-loop. Several important hydrophobic interactions to either side of the metal-binding loop are conserved in EF-hands [20]. These contacts are evident in crystal structures of EF-hand proteins such as calmodulin [18], on which this design is based. Some of these interactions (−4/+13; −5/+17) are inter-motif contacts, which promote dimerization when the EF-hand itself is well folded (numbers are relative to position +1 of the metal-binding loop, or D_{10} for P4). Other intra-motif hydrophobic contacts assist in stabilizing the Ca-binding pocket (−1/+13; −1/+16; +13/+17). These latter interactions, and in particular the −1/+13 contact, are apparently crucial in folding an EF-hand peptide into a rigid domain. Peptide P4 retains a hydrophobic contact at position −1/+16 (F_9–F_{25}), but the important hydrophobic pair at positions −1/+13 is absent (F_9–K_{22}). Thus, while peptide P4 has an EF-hand binding pocket, the fold may be expected to be less stabilized than a native EF-hand.

P4 NH_3-T_1ERRRFRVFD$_{10}$KDGNGYISAA$_{20}$EKIWFQNKRA$_{30}$KIK-COOH
P4a NH_3-T_1ERRRFDKDG$_{10}$NGYISAAELR$_{20}$HVKIWFQNKR$_{30}$AKIK-COOH
P3 NH_3-T_1ERRRQQLDK$_{10}$DGDGTIDERE$_{20}$IKIHFQNKRA$_{30}$KIK-COOH

Scheme 1.

The affinity of P4 for Eu(III) and Ce(IV) was studied by fluorescence titration, observing the change in W_{24} intensity as a function of added metal. The decrease in intensity due to subtle structural changes at W_{24} was well-fit to a 1:1 association model (Fig. 2), consistent with the lanthanide ion binding in the EF-hand loop. Additional equilibria,

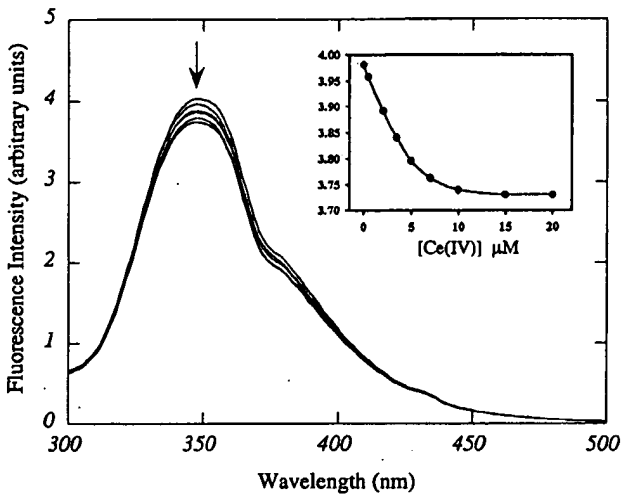

Fig. 2. P4 tryptophan fluorescence titration as a function of added $(NH_4)_2Ce(NO_3)_6$. Inset: intensity at λ_{max} vs. [Ce(IV)], and non-linear least squares fit to a 1:1 association model.

such as peptide dimerization or a second metal-binding event either do not occur or are not reported by changes in W_{24} fluorescence. The binding affinity of EuP4 was found to be log K_a=4.85, and for CeP4, log K_a=5.23. This Eu(III) affinity is weaker than that observed for the similar chimeric peptides P3 and P4a (EuP3, log K_a=5.0; EuP4a, log K_a=5.5) [11], and may reflect the relative spatial arrangement of the engrailed helices (varying the length of the EF-hand sequence results in the translation of the engrailed sequences along the helical axes) and loss of native inter-helical van der Waals and hydrogen bonding contacts. The weaker binding of P4 suggests that the metal affinity is tunable in this chimeric system as it is in native EF-hands, and emphasizes the importance of the −1/+13 stabilizing contact in designing robust peptide folds.

The peptide's affinity for Ce(IV) is greater than for the larger Eu(III) ion, though the latter more closely mimics the size of Ca(II). This suggests that P4 does not exhibit dramatic size discrimination, and the increase in affinity reflects the charge difference between the two ions. An adaptable coordination geometry is consistent with the flexible solution structure observed for these metallopeptides.

The binding of metals observed by fluorescence results in small but consistent changes in the CD spectrum at 50 μM P4 (Fig. 3). In the presence of equimolar Eu(III), the increase in negative intensity ($[\Theta_{222}]$, millidegrees) in the near UV is indicative of enhanced secondary structure. However, shorter helices are expected to contribute less to $[\Theta_{222}]$ than extended helices [23]. Additionally, the intensity is sensitive to fluxations in secondary structure, because the CD spectra reflect the time-averaged structure of the peptide [23]. Micromolar-metal affinity coupled with small apparent structural changes upon metal binding is con-

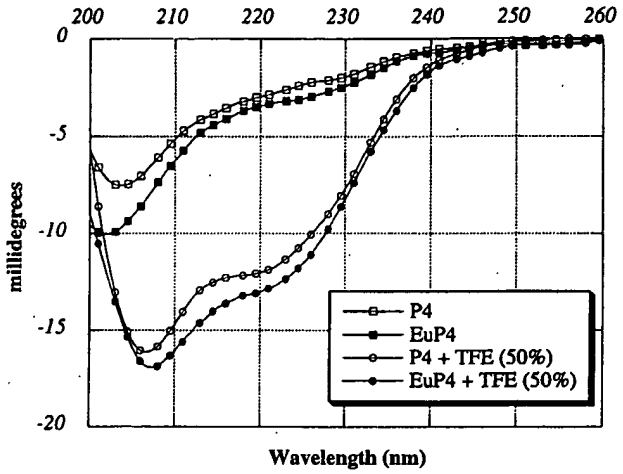

Fig. 3. Circular dichroism spectra of P4 (50 μM, 50 mM Tris buffer, pH 7.0) in the presence and absence of metals and TFE solvent.

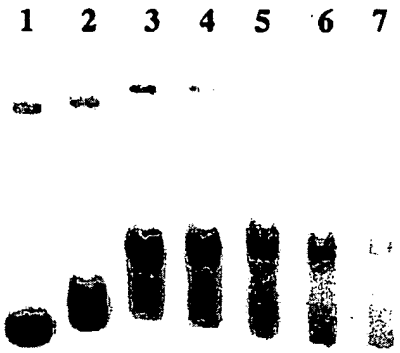

Fig. 5. Concentration dependent gel shift of pBR322 plasmid in the presence of increasing [EuP4]. Lane 1, plasmid control. Lanes 2–7, increasing EuP4 concentration: 10, 20, 30, 40, 50, 60 μM.

sistent with a very flexible, though not random, structure. When Eu(III) binds, it does not significantly enhance the rigidity of this fold.

When trifluoroethanol (TFE) is added, the amount of secondary structure observed in the CD spectrum increases for both free and metallated peptide. TFE is a solvent known to promote α-helix formation, and thus enhance secondary structure in proteins [24]. The shape of the curves suggests helical character, though not exclusively α-helical structure. The relative intensity of the peaks at 208 and 222 nm are not equal, which would be indicative of a purely α-helical structure imposed by the solvent. In the presence of TFE, the metal-binding event (Fig. 3, open to filled circles) organizes the secondary structure of the peptide more substantially than in aqueous buffer (open to filled squares). This again supports the conclusion that the metallopeptide has inherent structure, but is fluxional in solution.

Despite the flexibility of the chimeric metallopeptide, P4 is hydrolytically active in the lanthanide-bound form, catalyzing the cleavage of supercoiled DNA more efficiently than free Eu(III) or Ce(IV) ions. As a test of reactivity, peptide P4 was incubated with supercoiled pBR322 at 37 °C for 24 h in the presence and absence of Ce(IV) and Eu(III) (Fig. 4). The cleavage of supercoiled DNA (~50 μM b.p.) was followed by agarose gel electrophoresis, chelating the positively charged peptide immediately prior to analysis. A strong gel-shift would otherwise obscure products (Fig. 5).

As expected, P4 is completely inactive as a nuclease in the absence of metals. Additionally, at these low concentrations (10–25 μM), the free metal ions are not very reactive, with no observable cutting over control in 24 h. At pH 7, and in the presence of a weakly chelating buffer such as Tris, the lanthanides alone do not precipitate. At higher concentrations (>50 μM), significant nicking is observed for the free ions (data not shown).

In the presence of both P4 and the lanthanide ions Eu(III) and Ce(IV), however, the cleavage of supercoiled DNA is significantly enhanced. Comparatively efficient cleavage is observed with metallated P4 at concentrations

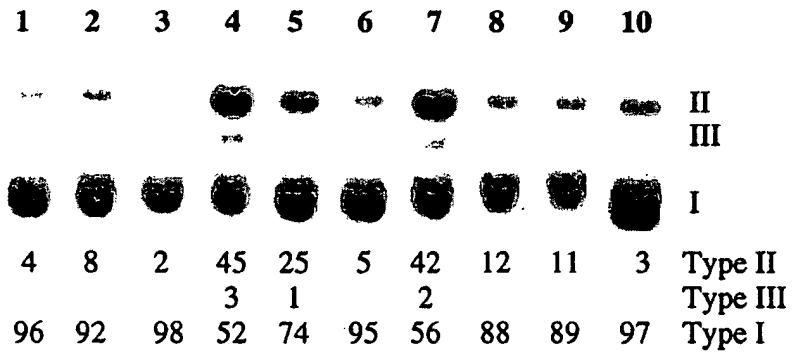

Fig. 4. Agarose gel electrophoresis assay of pBR322 DNA with various catalysts (50 mM Tris, pH 7.0). Quantified intensities of open circular, linear, and supercoiled bands are given below (as % of total DNA). Lanes 1 and 10: pBR322 plasmid control. Lane 2: 25 μM P4 (free). Lane 3: 25 μM $EuCl_3$. Lane 4: 10 μM EuP4 (1:1). Lane 5: 25 μM EuP4 (1:1). Lane 6: 25 μM $(NH_4)_2Ce(NO_3)_6$. Lane 7: 10 μM CeP4 (1:1). Lane 8: 25 μM CeP4 (1:1). Lane 9: 50:25 μM CeP4 (2:1).

near the thermodynamic dissociation constants. After 24 h, 45% of the total DNA was converted to open circular DNA (Form II) by 10 μM EuP4, and 42% by 10 μM CeP4 (Fig. 4, lanes 4 and 7). The cleavage is extensive enough to produce a small amount of linearized DNA (Form III). We are investigating whether this product is the result of double strand cleavage, rather than due to over-cleavage producing adjacent single strand cuts [6].

As the concentration of metallopeptide is raised, however, cleavage efficiency drops for both metals. This effect is particularly dramatic for Ce(IV) cleavage, for which only 12% cutting is observed by 25 μM metallopeptide. We have considered several explanations for this observation, including peptide dimerization, slow DNA-metallopeptide dissociation or precipitation, and metal-metal bridged dimers within the EF-hand pocket in the presence of excess free metal (below K_d). There is no evidence for metal-induced dimerization of P4 peptide, at least with the spectroscopic reporter of changes in tryptophan fluorescence. However, dimerization changes that do not further change the tryptophan environment would be transparent to this assay. Additionally, the metal-peptide binding equilibrium does not take into consideration stabilization of both the metal binding pocket and peptide dimerization interface that may occur in the presence of DNA. Thus, this explanation for the drop in cleavage efficiency cannot yet be dismissed.

It has been suggested that the outstanding reactivity of Ce toward DNA hydrolysis is in part dependent on the cooperativity of bridged Ce(IV)-hydroxide species [6,25], and it is tempting to consider such a structure for these exposed EF-hands. However, the addition of excess metal (Fig. 4, lane 9) failed to restore the reactivity seen at 10 μM CeP4, suggesting that 2:1 Ce:P4 complexes are not the active species.

Slow dissociation rates may decrease cleavage efficiency as the concentration of metallopeptide is raised. Gel shift assays show that the plasmid is fully bound in the presence of 25 μM metal-peptide, effectively lowering the concentration of uncomplexed DNA accessible for cleavage (Fig. 5). For CeP4 reactions, total DNA intensity is lost in cleavage gels at greater than 40 μM 1:1 Ce:P4 (data not shown). As no other products are observed, this intensity loss is attributed to the aggregation of the DNA imposed by the folded metallopeptide prior to electrophoresis, despite exposure to the cationic chelator resin. In fact, significantly less total DNA is present in lanes if samples are centrifuged prior to the addition of cation resin. This suggests that aggregation during the reaction effectively lowers both DNA and peptide concentrations, but removal of the peptide resolubilizes DNA.

The hydrolytic activity of EuP4 was also tested on a bisphosphate ester substrate other than DNA. The hydrolysis rates of the activated, small molecule substrate bisnitrophenyl phosphate (BNPP) were determined by following the formation of yellow nitrophenylate product by visible spectroscopy over the initial hours of the reaction. As shown in Table 1, the observed pseudo-first order rate constants are modest, but are approximately five orders of magnitude faster than uncatalyzed reaction under similar conditions (6×10^{-11} s^{-1}) [26]. Importantly, hydrolysis is faster in the presence of Eu-peptide than with free metal, as was seen with DNA substrate. It is notable that the rates level off at higher concentrations for both DNA and BNPP substrates, as reported earlier for the similar peptide P3 [11]. These data suggest that in addition to DNA precipitation, effects such as slow product dissociation or peptide dimerization may also contribute to the slower observed rates at high concentrations.

Table 1
Pseudo First-Order Rates of BNPP Cleavage as a Function of EuP4 at 37 °C

[Eu] (μM $EuCl_3$)	5	10	20	25	50	10
[P4] (μM peptide)	5	10	20	50	100	0
Rate (k_{obs}) ($s^{-1}\times10^7$)	0.8	3.9	4.3	5.7	5.9	1.4

The cleavage enhancements observed with these chimeric EF-hand peptides may arise from the cooperative function of peptide side-chains as general bases, and the efficient delivery of the active metal to the substrate by the positively charged sequences derived from engrailed. The enhanced reactivity at low concentrations may represent the optimum balance between metal-peptide K_a and tight metallopeptide–DNA interactions, which could slow nuclease dissociation rates and decrease total solution concentrations. It is clear that the synergy of free and peptide-bound metal results in significantly faster cleavage rates, and the mechanism and character of this association is currently being investigated.

4. Conclusions

The ability to cleave DNA has now been demonstrated for HTH/EF-hand chimeric peptides with significant sequence variations, implying that the catalytic potential of these systems is relatively universal [10,11]. For the chimeric peptide P4, the EF-hand motif binds Eu(III) and Ce(IV) ions, and this binding event enhances the structure of the peptide. However, it does not result in a rigid, well-defined peptide structure, and thus the lessons learned in this design will inform future chimera sequences. The binding affinity and flexibility of the peptides may be tunable, and retaining the relative register of the HTH helices and the −1/+13 hydrophobic contacts appears crucial to generating a static fold. Despite this flexibility, the metallopeptide is an effective catalyst, capable of significantly enhancing the cleavage of DNA. In contrast,

at these concentrations neither the free peptide nor the free metal ions are effective catalysts for DNA cleavage. This suggests that the sequences derived from the HTH motif, containing many positively charged lysine and arginine residues are in effect delivering the metal to the DNA backbone for cleavage. While a rigid structure may be preferable for the design of selectivity, it is not a criterion for reactivity.

Acknowledgements

We thank the organizers of ICBIC 10, Florence, Italy, for a stimulating meeting, and the reviewers for many useful comments. This work was supported by NSF CAREER grant 0093000.

References

[1] W.F. DeGrado, C.M. Summa, V. Pavone, F. Nastri, A. Lombardi, Annu. Rev. Biochem. 68 (1999) 779–819.
[2] G. Xing, V.J. DeRose, Curr. Opin. Chem. Biol. 5 (2001) 196–200.
[3] Y. Lu, S.M. Berry, T.D. Pfister, Chem. Rev. 101 (2001) 3047–3080.
[4] A.V. Efimov, FEBS Lett. 391 (1996) 167–170.
[5] S.J. Franklin, Curr. Opin. Chem. Biol. 5 (2001) 201–208.
[6] M.E. Branum, A.K. Tipton, S. Zhu, L. Que Jr., J. Am. Chem. Soc. 123 (2001) 1898–1904.
[7] L.L. Chappell, D.A. Voss, W.D. Horrocks Jr., J.R. Morrow, Inorg. Chem. 37 (1998) 3989–3998.
[8] A. Roigk, R. Hettich, H.-J. Schneider, Inorg. Chem. 37 (1998) 751–756.
[9] M. Komiyama, Chem. Commun. 16 (1999) 1443–1451.
[10] Y. Kim, J.T. Welch, K.M. Lindstrom, S.J. Franklin, J. Biol. Inorg. Chem. 6 (2001) 173–181.
[11] J.T. Welch, M. Sirish, K.M. Lindstrom, S.J. Franklin, Inorg. Chem. 40 (2001) 1982–1984.
[12] C.S. Wilcox, in: H.-J. Schneider, H. Durr (Eds.), Frontiers in Supramolecular Organic Chemistry and Photochemistry, VCH, Weinheim, 1991, pp. 125–139.
[13] M. Sirish, H.-J. Schneider, Chem. Commun. 10 (1999) 907–908.
[14] J.-M. Pfefferle, J.-C.G. Bünzli, Helv. Chem. Acta 72 (1989) 1487–1494.
[15] Y.-H. Chen, J.T. Yang, H.M. Martinez, Biochemistry 11 (1972) 4120–4131.
[16] R. Hettich, H.-J. Schneider, J. Am. Chem. Soc. 119 (1997) 5638–5647.
[17] R.P. Hertzberg, P.B. Dervan, J. Am. Chem. Soc. 104 (1982) 313–315.
[18] R. Chattopadhyaya, W.E. Meador, A.R. Means, F.A. Quiocho, J. Mol. Biol. 228 (1992) 1177.
[19] L. Tucker-Kellogg, M.A. Rould, K.A. Chambers, S.E. Ades, R.T. Sauer, C.O. Pabo, Structure (London) 5 (1997) 1047–1054.
[20] J.J. Falke, S.K. Drake, A.L. Hazard, O.B. Peersen, Q. Rev. Biophys. 27 (1994) 219–290.
[21] M.R. Nelson, W.J. Chazin, Biometals 11 (1998) 297–318.
[22] A. Lewit-Bentley, S. Réty, Curr. Opin. Struct. Biol. 10 (2000) 637–643.
[23] J.D. Hirst, C.L.I. Brooks, J. Mol. Biol. 242 (1994) 173–178.
[24] P. Luo, R.L. Baldwin, Biochemistry 36 (1997) 8413–8421.
[25] R. Ott, R. Krämer, Appl. Microbiol. Biotechnol. 52 (1999) 761–767.
[26] J. Chin, M. Banaszczyk, V. Jubian, X. Zou, J. Am. Chem. Soc. 111 (1989) 186–190.